

# Smart Generation System of Personalized Advertising Copy and Its Application to Advertising Practice and Research

Deng, Shasha; Tan, Chee Wee; Wang, Weijun; Pan, Yu

*Document Version*  
Accepted author manuscript

*Published in:*  
Journal of Advertising

*DOI:*  
[10.1080/00913367.2019.1652121](https://doi.org/10.1080/00913367.2019.1652121)

*Publication date:*  
2019

*License*  
Unspecified

*Citation for published version (APA):*  
Deng, S., Tan, C. W., Wang, W., & Pan, Y. (2019). Smart Generation System of Personalized Advertising Copy and Its Application to Advertising Practice and Research. *Journal of Advertising*, 48(4), 356-365.  
<https://doi.org/10.1080/00913367.2019.1652121>

[Link to publication in CBS Research Portal](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

## Take down policy

If you believe that this document breaches copyright please contact us ([research.lib@cbs.dk](mailto:research.lib@cbs.dk)) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 18. Jun. 2025



# Smart Generation System of Personalized Advertising Copy and Its Application to Advertising Practice and Research

**Shasha Deng, Chee Wee Tan, Weijun Wang, and Yu Pan**

Journal article (Accepted manuscript\*)

**Please cite this article as:**

Deng, S., Tan, C. W., Wang, W., & Pan, Y. (2019). Smart Generation System of Personalized Advertising Copy and Its Application to Advertising Practice and Research. *Journal of Advertising*, 48(4), 356-365.  
<https://doi.org/10.1080/00913367.2019.1652121>

This is an Accepted Manuscript of an article published by Taylor & Francis in *Journal of Advertising* on 21 Aug 2019, available online:

DOI: <http://www.tandfonline.com/10.1080/00913367.2019.1652121>

\* This version of the article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the publisher's final version AKA Version of Record.

Uploaded to [CBS Research Portal](#): August 2020

# **Smart Generation of Personalized Advertising Copy and Its Application to Advertising Practice and Research**

**Shasha Deng**

School of Business and Management  
Shanghai International Studies University  
550 Dalian Road, Shanghai, 200083, China  
Email: [shasha.deng@shisu.edu.cn](mailto:shasha.deng@shisu.edu.cn)

**Chee-Wee Tan**

Department of Digitization  
Copenhagen Business School  
Howitzvej 60, 5.24, 2000 Frederiksberg, Denmark  
Email: [ct.digi@cbs.dk](mailto:ct.digi@cbs.dk)

**Weijun Wang\***

Key Laboratory of Adolescent Cyber-psychology and Behavior  
School of Psychology  
Central China Normal University  
152 Luoyu Road, Wuhan 430079, Hubei, China  
Email: [wangwj@mail.ccnu.edu.cn](mailto:wangwj@mail.ccnu.edu.cn)

**Yu Pan**

Laboratory of Applied Brain and Cognitive Sciences  
School of Business and Management  
Shanghai International Studies University  
550 Dalian Road, Shanghai 200083, China  
Email: [13311887777@163.com](mailto:13311887777@163.com)

*\*Corresponding Author*

Paper Accepted at *Journal of Advertising* (JA).  
Please do not quote or cite without permission. Comments are welcome.  
Please address all correspondence to the first author.

# Smart Generation of Personalized Advertising Copy and Its Application to Advertising Practice and Research

## *Abstract*

*Artificial intelligence in programmatic advertising constitutes fertile grounds for marketing communication with tremendous opportunities. Yet, despite its touted benefits, contemporary implementations of programmatic advertising do not harness self-generative technologies so much so that different consumers are exposed to identical content. Consequently, we advance a Smart Generation System of Personalized Advertising Copy (SGS-PAC) that can automatically personalize advertising content to align with the needs of individual consumers. Analytical results from a user experiment involving about 80 subjects underscore that personalized advertising copies generated by SGS-PAC can bolster click rate in online advertising platforms. Findings from this study bear significant implications for the application of artificial intelligence in online advertising.*

**Keywords:** Advertising copy personalization, text analytics, social media, natural language processing, programmatic advertising

Advertising is an indispensable form of publicity for businesses. With disappearing media boundaries, artificial intelligence and big data analytical techniques have become the main driving force in reorganizing information flows and streamlining operational processes that underpin advertising (Malthouse and Li 2017, Kietzmann, Paschen and Treen 2018). In turn, conventional advertiser-centric advertising has given way to consumer-centric programmatic advertising (Campbell, et al. 2011).

Conventional advertising accentuates ‘media buying’ in that it place emphasis on purchasing advertising space from media providers through daily subscriptions or via Cost per Mille (CPM) arrangements (Kumar and Gupta 2016). Consequently, advertising content broadcasted across diverse media, at different time periods, and on separate occasions seldom differ in conventional advertising. In contrast, programmatic advertising concentrates on ‘audience purchase’ by aligning advertising content with consumer attributes (Delbaere, McQuarrie and Phillips 2011). In this sense, programmatic advertising denotes an advertising cycle that blends programmatic ideas with consumer insights and impression purchase. It places emphasis on programmatic technology as the base with latent multi-screen data access and multidimensional big data analytics as the core. It also incorporates the RTB (Real Time Bidding) ecosystem, Mobile DSP (Demand Side Platform), and multi-screen programmatic buying as supporting architectures (Schultz 2016). Consequently, programmatic advertising accommodates the demands of individual consumers by displaying personalized advertising content that resonates with each consumer. This implies that consumers, who embrace the same brand but exhibit dissimilar attributes, will be presented with distinct advertising content at different time periods and on separate occasions. Coupled with the availability of big data,

programmatic advertising is transforming the advertising industry by yielding consumer insights that foster creativity in the production of advertisements (Malthouse and Li 2017).

Yet, despite its touted benefits, the realization of programmatic advertising comes with a major challenge. A huge pool of copywriters is required to generate multiple copies of the same advertisement to cater to individual consumers as envisioned in programmatic advertising and even then, they may not be adequate. This is because one cannot rely on copywriters alone to relentlessly churn out fresh advertising content in a timely fashion. To tailor advertising content to individual consumers, copywriters have to process massive amounts of information in terms of comprehending the needs of these consumers. In turn, this becomes a strain on efficiency whenever the exponential increase in information exceeds copywriters' processing capability (Abbasi, et al. 2018). Conversely, automating the advertisement production process by merely repeating the brand or replacing the product name, does not work well because it will undoubtedly lead to content standardization. Without the aid of self-generative technologies, copywriters can easily get lost in an avalanche of information, a phenomenon known as *information overload*.

To address the abovementioned knowledge gap within extant literature, we draw on advanced text analytical techniques that are capable of advertising copy to match the needs of individual consumers. Particularly, we devise a Framework for Smart Generation System of Personalized Advertising Copy (SGS-PAC) to support the analysis and generation of personalized advertising copies. Results from a user experiment attest to the augmented personalization capabilities afforded by SGS-PAC, which in turn bear significant implications for the deployment of artificial intelligence in advertising.

The remainder of the paper is organized as follows. We review previous work relating to online advertising, personalization, advertising copy, and automated text summarization in the next section. Following which, we describe, in Section 3, our SGS-PAC framework that maps out a system for automatically generating personalized advertising copy through combining advanced text mining and machine learning techniques. Grounded in the framework, we further illustrate, in Section 4, the utility of such automated generative process in advertising across various product scenarios before we conclude with the implications of our findings for theory and practice in Section 5.

## **THEORETICAL FOUNDATION**

### **Online Advertising and Personalization**

The emergence of Web 3.0 has radically altered the manner by which individuals solicit and consume information in fulfillment of their personalized informational needs. Recognizing the value of personalization, businesses have endeavored to deliver information services that are much more intelligent and precise in accommodating individual needs (Liu, Burns and Hou 2017). Particularly, marketers are operating in increasingly digitalized environments with expanded avenues for reaching out to consumers on a personal basis (Becker, Linzmajer and von Wangenheim 2017, Schultz 2016, Kumar and Gupta 2016). This in turn opens up opportunities for marketers to engage in innovation within the online advertising space, especially with regards to the realization of personalized advertising. Personalized advertising is founded on an intricate understanding of consumer behavior and preferences. It relies on the collection and analysis of pertinent behavioral data — in the likes of blogs, comments, likes,

searches, and tweets — to draw inferences about consumer preferences and generate personalized advertising content (Kietzmann, Paschen and Treen 2018). Personalized advertising thus departs from conventional approaches that allow consumers to articulate their own preferences even though these stated preferences have been discovered to be rather arbitrary and may not be reflective of the former's actual needs (Zenil, et al. 2014). Indeed, past studies have demonstrated that the endorsement effect of advertisements is stronger when advertising content is aligned with the profile of its target audience (Simsek and Karagoz 2018).

### **Advertising Copy and Automated Text Summarization**

Advertising is most effective when the type of content embodied in the advertising copy matches the information that consumers are seeking at the moment, rendering the message easier to process and recall (Baker and Lutz 2000). For this reason, online behavioral traces of consumers constitute a critical source of data on their informational needs and is indispensable in the delivery of personalized advertising content (Liu-Thompkins and Malthouse 2017).

With advances in natural language processing, automatic summarization has emerged as a prevalent method for extracting focal sentences from a collection of texts. Based on techniques for comparing sentence similarity, automatic summarization can identify focal sentences through sentence scoring and discrimination. When redundant sentences are removed, the remaining sentences are reordered (Dief, et al. 2017). To bolster sentence similarity and enhance text readability, prior research has advocated a novel summarization technique that incorporates cuckoo search to solve the problem of sentence selection in multi-text summarization (Rautray and Balabantaray 2018). Yet, despite the merits of automatic summarization, there is a paucity of studies that have applied automatic text generation in

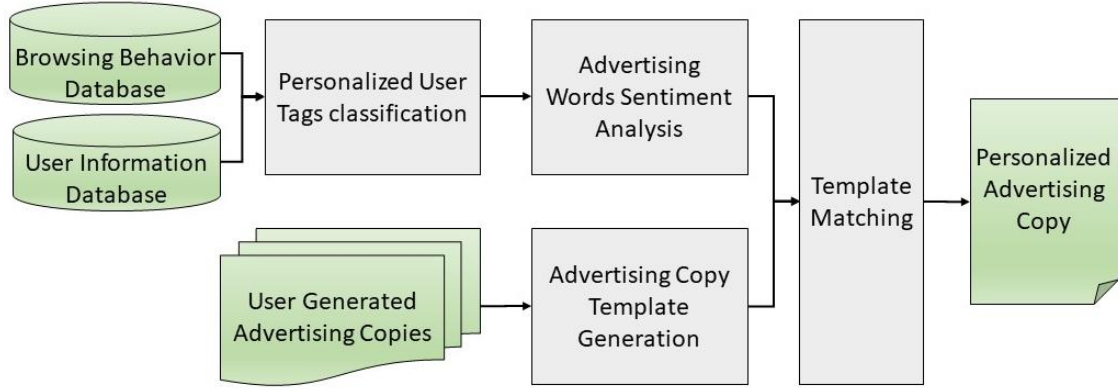


advertising research. Given that consumers are bombarded with massive amounts of information on a daily basis, it would greatly improve their receptivity toward advertising copies if the content of these messages can be personalized according to behavioral data.

To this end, we synthesize algorithms for non-negative matrix decomposition and semantic analysis to derive multiple objective functions for assessing sentence similarity. Combined with word frequency statistics, this study strives to attain both diversity and relevance in the area of text summary (Ansamma, Premjith and Wilscy 2017).

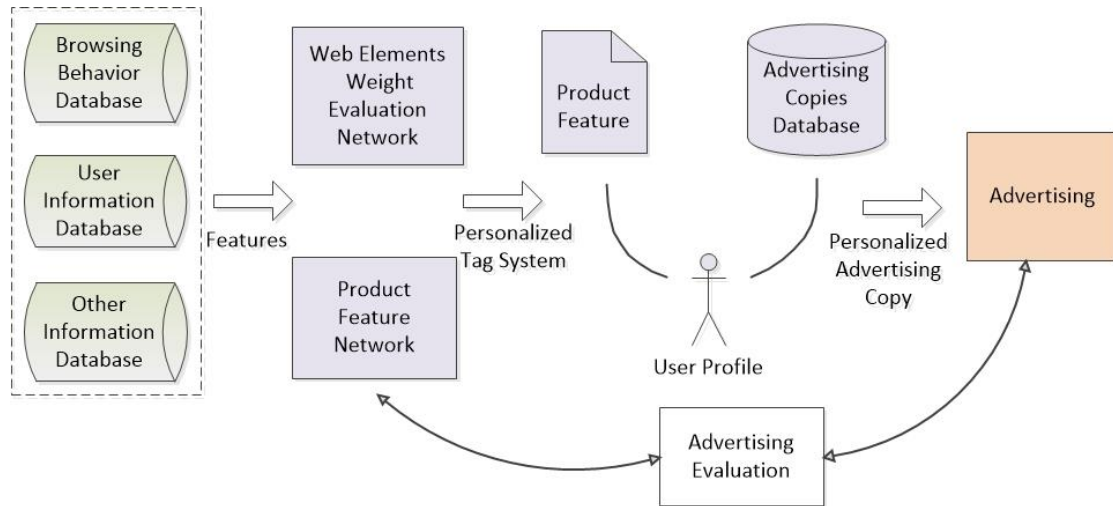
## **TOWARD A SMART GENERATION SYSTEM FRAMEWORK OF PERSONALIZED ADVERTISING COPY**

We have devised a Smart Generation System of Personalized Advertising Copy (SGS-PAC) that can automatically generate personalized advertising copies to match the demands of programmatic advertising (see Figure 1). The system comprises four key components, namely personalized user tag classification, advertising words sentiment analysis, advertising copy template generation, and template matching. For each consumer, the system will produce a specific advertising copy, which aims at the characteristics of the current user and is connected with the scene the user situates at the moment. In the following, the technical implementation of these four components will be briefly discussed.



**FIGURE 1. Framework for Smart Generation System of Personalized Advertising Copy (SGS-PAC)**

The SGS-PAC is dependent on the availability of the massive amount of data on user search behavior, user information, product information, etc. Its specific process of generating the personalized advertising copies is shown in Figure 2. First, data modeling and neural network methods are used for acquiring the related features to develop the web elements weight evaluation network and product feature network. Resting on these two network models, a personalized tag system can be established. Then, the tag words in the system are analyzed for their affective states, and the words with positive emotional polarity are selected as the product features and user features. The fourth step is to study the existing advertisements to generate corresponding advertisement templates. A personalized advertisement copy for the user is generated by matching the user's characteristics, product features, and the advertisement template, to be then delivered to the user. Finally, the parameters are adjusted by evaluating the effectiveness of the system.



**FIGURE 2. Illustration of SGS-PAC process**

### **Personalized User Tags Classification Based on Web Elements Weight Evaluation and BP Neural Network**

In order to provide users with personalized ads, we need to clearly understand users' personal preferences for different categories of products. Generally speaking, advertising companies will build different user tags system for different product categories. For example, in the tags system of maternal and infant products, the tags categories include pregnancy stage, parenting stage, mother focus, baby focus and so on. Among them, the tags of "mother focus" category include postpartum weight loss, parent-child life, second child, maternity, postpartum health and so on. Therefore, the research framework needs to select appropriate tags in each tag category for each user according to user preferences.

The exiting psychological studies have proved that people's perception, understanding and recall are affected by what they have seen before (Bargh 2006). Therefore, according to the browsing behavior of users and the history of shopping record, we realize the user tags automatic classification.

### ***Calculation of Web Elements Weights through Modified Analytic Hierarchy Process (AHP)***

People tend to be only interested in a part of the web for a long time, and this part of the web is known as the "hot zone". Existing studies provide many measures including title, comment area content, comment area label, picture, browsing time effect, etc. However, the specific index weight value is not given. The improved AHP can be adopted to identify index weights (Saaty 2008). The method is often used in research selection and project evaluation decision in previous studies (Velasquez and Hester 2013, Guariguata, et al. 2014). In order to avoid preferences of a small number of raters lead to the index weight calculation unreasonable, the method will be the first to calculate weight of raters evaluation, then uses the traditional AHP method to calculate weight indicators at various levels, lastly computing comprehensive weight of each indicator through the weighted average between all level index weights and rater evaluation weight.

The process of identifying the web elements weights on the basis of the modified Analytic Hierarchy Process (AHP) can be divided into the following steps:

#### **① Calculating the weight of rater evaluation**

The multi-index evaluation system is composed of  $n$  raters  $o_1, o_2, \dots, o_n$  and  $m$  indicators  $I_1, I_2, \dots, I_m$ , where  $\chi_{ij} = I_j(o_i)$  ( $i = 1, 2, \dots, n, j = 1, 2, \dots, m$ ) represents the rater  $o_i$ 's score for the index  $I_j$ . The evaluation data matrix can be expressed as  $A = [\chi_{ij}]_{n \times m}$

The correlation coefficient of the raters was calculated through the raters' evaluation data matrix. Then the evaluation matrix of raters is calculated based on the planning model of raters to maximize their own advantages. The comprehensive evaluation values of all raters are

obtained according to the evaluation matrix. Finally, the weight of each rater in the rater group was determined.

② Calculating all level index weights using traditional AHP method

The traditional AHP method used pair-wise comparison to construct the comparison matrix. In order to make the rater's judgment quantified, the 1-9 scale method was used for measurement. Finally, develop the weights for all level index and criteria by calculating and checking the Consistency Ratio(Saaty 2008).

③ Computing Comprehensive Weight of Each Indicator

The weight of the first-level indicator is expressed as  $U_j = \sum_{i=1}^n \lambda_i \omega_{ij}$ , where  $\omega_{ij}$  means the weight of the rater  $i$ 's first-level indicator  $j$  that is calculated by the traditional AHP method.  $\lambda_i$  presents the weight of the rater  $i$  in the rater group. The weight of the second-level indicator is written as  $U_{jk} = \sum_{i=1}^n \lambda_i \omega_{ijk}$ , where  $\omega_{ijk}$  represents the weight of the rater  $i$ 's second-level indicator  $jk$ .

***Predict User Tags Using Back Propagation (BP) Neural Network Model***

Back Propagation Neural Network (BP NN) model is a common method used in combination with optimization strategy (such as gradient descent method) to train artificial neural network (Hecht-Nielsen 1992). The Word2Vec tool is used to construct word vectors for the content of web pages. Combined with the web element weights obtained previously and user browsing habits, the value of each user's word vector can be accumulatively calculated the weight value of web page elements. The results of each user's word vectors can be used as the input of BP NN model. The output of the BP NN model is a series of tags in a certain tag category of the user. Take the "mother focus" tag category for example, a BP NN model is

used to predict one of tags (such as maternity) in the tag category for each user. Different BP NN models are used to predict user tag in different tag categories.

The users' tags in different tag categories will provided by advertising company. These uses' tags already were proven to be effective in click-through rate prediction for many times. During training of the BP NN models, these users' tags will be viewed as target value  $Y$ . According to user's browsing history and the weight of web element, a 300-dimensional words vector will be used as a set of features  $X = x_1, x_2, \dots, x_{300}$  for BP NN models. We import MLPClassifier from sklearn library to construct BP NN classifier, and use grid search method to select optimal parameters of each BP NN model.

### **Advertising Words Sentiment Analysis Incorporating Mood State Lexicon**

Product pages and user review content are full of emotive words that express feelings. Profile of Mood States (POMS) has been widely used for the assessment of mood states with short duration and specific situation which has significant impact on customer's judgement and decision (McNair, Droppleman and Lorr 1992). This study first obtains the mind-state seed word set based on the POMS, a psychometric tool, and extends the POMS seed word set through the synonymy thesaurus and the Word2vec semantic similarity based on the large-scale microblog corpus (McNair, Droppleman and Lorr 1992, Bollen, Mao and Pepe 2011). Then it uses the Markov Random Walk algorithm to identify the mood state values of each dimension to automatically construct a multi-dimensional, fine-grained emotional state dictionary.

### ***Network Construction of Mood State Lexicon***

The way to construct the mood state lexicon network is to first select the mood words or phrases in the vocabulary system of the Mental State Scale, and then to obtain 105 seed words with the same meaning based on the synonymy thesaurus - Tongyici Cilin (Parkinson, et al. 1996, Jiu-le and Wei 2010).

First, based our experience, the similarity threshold limit value is set to 0.8. Higher than 0.8 will lead to less synonyms for the seed words, while lower than 0.8 will lead to some irrelevant words being selected. The vocabulary with the value greater than the threshold value is obtained as the candidate word, computed by semantic similarity algorithm in Tongyici Cilin (Jiu-le and Wei 2010).

Second, the word embedded tool—Word2Vec is used to construct a word vector on the corpus, and the seed words defined as the vocabulary with the correlation greater than the threshold value, iterate to obtain the candidate words;

Third, the mind state seed words are associated with the candidate words based on the similarity degree to establish a semantic network of mood state Chinese vocabulary.

### ***Calculation of Vocabulary Mood State Dimension Based on Random Walk***

Based on the Synonymy Thesaurus and Word2vec lexical similarity measure, any pair of semantically related candidate words are connected to form a mood vocabulary semantic network graph  $G(W, E)$ .  $W$  represents all vocabulary sets.  $E$  means an edge set connecting pairs of vertices of all the words (a set of semantic similarities between two words). In the lexical semantic network graph  $G$ , with the seed words as the absorbing boundary conditions,

the lexical multi-dimensional scores are computed by the Markov Random Walk algorithm that proposed by Hassan and Radev (2010). Finally we obtain a fine-grained mood state dictionary with multiple mood state dimension values.

A random walk is a mathematical statistical model that consists of a series of randomly generated trajectories. First, in the semantic graph  $G$ , the transform probability  $P_{ij}$  represent the moving probability from node  $i$  with a unknown mood state value to node  $j$ . It will not stop until it reaches a node with known mood state value. Words with known mood state values in each dimension will serve as the absorption boundary of the random walk. Repeat  $N$  times, and the percentage of time required to reach known mood state word node in certain mood state category can be regarded as mood state value in the mind dimension of unknown node.

### **Automatic Generation of Advertising Copy Template**

Our goal is to automatically generate an advertising copy template by learning about the texts of the existing advertisement copies. The template is mainly made up of the advertising copy category, the copy template, and the candidate word chain corresponding to each template.

#### ***Advertising Copy Category***

The types of advertising copy can be divided into sales copy and communication copy based on its purpose. The aim of communication copy is to induce public recognition through the promotion, mainly to expand brand influence, establish brand image, and promote corporate culture, instead of increasing the sales directly with the aid of commercials. The sales copy is aimed to promote the sales through advertising to increase profit from direct sales.



According to the categorization of advertising copies, we defined four types of advertising copies: product, brand, word of mouth, and promotion. Two independent coders independently annotated the advertising copies, attaining a Cohen's Kappa value of 84.5% for inter-coder reliability. We extracted the feature set including: (1) binary/presence vector for all nouns and verbs appearing at least three times in the training corpus, lemmatized with their part-of-speech information; and (2) whether or not the message has sentiment. These features are input into a series of linear support vector machines (SVM) classifiers (Cristianini and Shawe-Taylor 2000). The output of a series of SVM is the purpose of advertising copy (sales copy and communication copy) and the type of advertising copy (product, brand, word of mouth, and promotion)

### ***Textual Analysis of Advertising Copy***

To perform content analysis, the syntactic structure of advertising copy's content needs to be identified. For English language, Stanford Parse is a popular linguistic parsing tool which uses Penn Treebank tag set. And similar tools exist for other languages such as Language Technology Platform for Chinese language. These tools perform word segmentation, Part-of-Speech (POS) tagging, named entity recognition (NER) and dependency tree parsing. Particularly, the ability to segment and identify noun, adjective, and idiom compounds will greatly improve information extraction accuracy.

Semantic role labeling referred to as latent semantic analysis, means analyzing the semantic relationship among content words in a sentence. Several kinds of initial templates of the advertising copy will be defined initially based on the semantic dependency in the text. The initial template will define the most basic and simplest grammar and lexical rules. The

dependency structure is first established according to the dependency grammar. After all the modifiers and core word pairs in the sentence are found, the semantic relationship will be identified for the word pairs. Finally, a new advertisement copy rule is constructed. If this advertising copy generation rule is not included in the template gallery, it will then be added.

### **Personalized Advertising Copy Generation Based on Template Matching**

After the advertisement copy template is selected, the tag in the personalized advertisement tag system will be matched with the candidate word chain in the template gallery to choose the corresponding word and advertisement rule.

### ***Similarity Measure between Keywords and Template Candidate Words***

This study combines the computation methods of both the semantics and word similarity to adapt to short advertising copies. Word shape similarity analysis compares the string similarity computed by Jaro-Winkler distance, which can better address the input errors and problems caused by the synonyms with similar shapes. Semantic similarity calculation introduces Word2vec and the *Synonymy Thesaurus* (extended version), using the nearest word neighbors in Word2vec as the feature, and measuring the distance of the phrases rather than the words for the improvement of the Minimum edit distance formula to calculate the semantic similarity, and eliminating the obstacles to the synonym similarity computation of the advertising copy mixed with both English and Chinese.

### ***Personalized Advertising Copy Based on Template Matching***

The first step is to select the type of advertisement that are adaptive to both the needs of the advertiser in the programmatic advertisement and the personalized tag of the user, and to choose its corresponding advertisement template.

Then the appropriate keyword will be selected to fill in the template in accordance with the degree of similarity between the keyword and the candidate word. If the template contains a function, the content of the final generated advertisement copy can be determined by the template's property settings. The function has three indicators: Count, Filter, and Compare. These three indicators can be used to correct the syntactic patterns in the advertisement copy and the word choice, including the selection of numbers, named entities, and adjectives. The related templates can also be merged based on the logic relations.

## **EMPIRICAL VALIDATION OF THE SGS-PAC**

System performance does not always correlate with user performance, especially for complex tasks (Turpin and Scholer 2006). To distinguish the effectiveness of advertising copy generated by SGS-PAC versus that written by copywriters in actual application scenarios, we conducted a within-subject laboratory experiment in which subjects were exposed to advertising copies generated by both SGS-PAC and copywriters before being asked to indicate which advertising copy they favor.

### **Experimental Design**

The experiment was based on real commercial data supplied by an external research partner who operates a third-party retail platform. Through our research partner, we were able

to gain access to consumers' browsing and purchase behaviors (e.g., attention span, adding products to shopping cart, and making purchases), product attributes, entrance IP address to the retail platform, and advertising content. From the commercial data, we were able to extract three categories of products and four types of consumer profiles. Specifically, we identify three categories of products that are clicked on or purchased most frequently through the retail platform: ①Mobile Phones, ②Outdoor Jacket, and ③Washing Machine. For each product category, we further chose a product corresponding to a household brand. We then classify consumers, who click to view or purchase these three products through the retail platform, into four types based on user tags that have been previously employed by advertising companies to segment their target audience. These four types of consumer profiles include: ①Type 1 consumers who focus on the appearance, quality, and workmanship of the product; ②Type 2 consumers who prioritize safety and value-for-money; ③Type 3 consumers who pay attention to functionality, operability, and response time, and; ④Type 4 consumers who prefer affordable and cost-effective products. From the interplay between product categories and consumer profiles, we arrived at twelve product-consumer pairs to be evaluated by experimental subjects.

For each product-consumer pair, two treatment conditions were introduced. In the first treatment condition, personalized advertising copies were generated through the SGS-PAC and presented to experimental subjects. In the second treatment condition, several advertising copies of the target product, which had been displayed at least once on multiple retail platforms such as Toutiao, were selected from an online advertising company. Each experimental subject was exposed to both treatment conditions for the twelve product-consumer pairs. The treatment

conditions were randomized to eliminate ordering bias. For each product-consumer pair, experimental subjects were instructed to answer a question about their preference for system-versus human-generated advertising copy: *“If you are a consumer who focuses on the appearance, quality, and workmanship of the product, which advertising copy of [a target product] do you favor?”*

### **Data Collection**

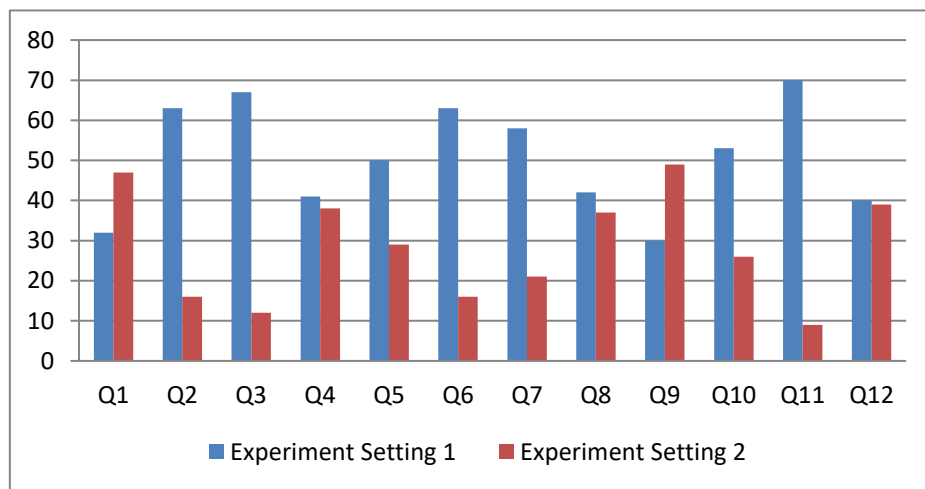
The experiment was pretested with two doctoral students before a pilot study was conducted with five doctoral and seven master’s students. Based on the feedback garnered from the pilot study, we further clarified the wording in the questions as well as refined the experimental instructions and procedures. For the actual experiment, subjects were recruited via email invitations to faculty members as well as Master of Business Administration (MBA) and graduate students from a large university in China. A total of 96 subjects participated in the actual experiment, none of whom had participated in the pilot study. Subjects who failed to answer all twelve questions or follow experimental instructions were excluded, yielding an eventual sample of 79 data instances for analysis.

### **Data Analysis**

Figure 3 depicts subjects’ selection of advertising copies across the two treatment conditions. Whereas the first treatment condition is denoted by the blue bar, the second treatment condition is represented by the red bar. Analytical results indicate a clear preference among subjects for the advertising copies generated by the SGS-PAC in Q2, Q3, Q4, Q5, Q7, Q10, and Q11. Conversely, subjects exhibit no obvious preference towards advertising copies

generated by both treatment conditions in Q6, Q8, and Q12. Finally, for Q1 and Q9, subjects seem to prefer advertisement copies written by copywriters.

Taking Q11 as an illustrative example, for Type 3 consumers who pay attention to functionality, operability and response time, the advertising copy generated through the SGS-PAC is: “LITTLE SWAN” washing machine, smart hand washing, energy saving and quieter, brings you a comfortable new life. On the other hand, the advertising copy selected from an online advertising company is: “LITTLE SWAN” intelligent washing machine, satisfy the mood of every dress. From above, it is apparent that the SGS-PAC would generate distinctive advertising copies for each product-consumer pair according to the profile of consumers. This stands in contrast to the advertising copies written by copywriters in that they typically contain literal descriptions of the product and try to establish a match with every consumer regardless of his/her profile.



**FIGURE 3. Overview of dataset**

Table 1 summarizes the analytical results of our user experiment. Paired *t*-tests were conducted to compare subjects’ preference for advertising copies between the two treatment conditions. Results indicate that the first treatment condition significantly outperforms the

second ( $p < 0.001$ ) across all twelve product-consumer pairs. This in turn lends credibility to the effectiveness of personalized advertising copy generated by SGS-PAC as a viable approach for promoting various products to diverse consumers.

**TABLE 1**

**Results across All 12 Questions for Human Subject Experiment**

<b>Experiment Setting</b>	<b>Average Number of Subjects Choice</b>
<b>1</b>	7.71*
<b>2</b>	4.29

\* Significantly outperformed Experimental Setting 2;  $p < 0.001$

Next, we delve into the impact of distinct product categories and consumer profiles on the latter's preference for system- versus human-generated advertising copies.

**TABLE 2**

**Results for Three Types of Product Questions**

<b>Experiment Setting</b>	<b>Mobile Phones</b>	<b>Technical Jacket</b>	<b>Washing Machine</b>
<b>1</b>	2.57*	2.70*	2.44*
<b>2</b>	1.43	1.30	1.56

\* Significantly outperformed Experimental Setting 2;  $p < 0.001$

Analytical results for the three product categories across the two treatment conditions are summarized in Table 2. Consistent with our aggregated findings, the first treatment condition significantly outperforms the second across the three product categories.

**TABLE 3**

**Results for Four Types of Customer Role Questions**

Experiment Setting	Type 1	Type 2	Type 3	Type 4
1	1.42 <sup>#</sup>	2.27 <sup>*</sup>	2.47 <sup>*</sup>	1.56 <sup>+</sup>
2	1.58	0.73	0.53	1.44

\* Significantly outperformed Experimental Setting 2;  $p < 0.001$

# Did not significantly outperform Experimental Setting 2;  $p = 0.419$

+ Did not significantly outperform Experimental Setting 2;  $p = 0.535$

Table 3 summarizes the analytical results for the four consumer profiles. As anticipated, for Type 2 and 3 consumers, the first treatment condition significantly outperforms the second (all pair-wise  $t$ -test  $p$ -values  $< 0.001$ ). For Type 1 and 4 consumers, the performance gain of system-generated advertising copies over their human-generated counterparts was insignificant ( $p = 0.419$ ;  $p = 0.535$ ). These results indicate that personalizing advertising copies to suit Type 2 and 3 consumers might not yield tangible benefits as compared those written by copywriters. Moreover, Type 1 and 4 consumers who opted for personalized advertising copies generated by the SGS-PAC did not significantly outperform those written by copywriters, implying that the former may not be overly effective for the generation of advertising copy for such consumers. All in all, analytical results presented in Tables 1, 2 and 3 support our contention that personalized advertising copies generated by SGS-PAC could facilitate enhanced user click rate in online advertising.

## CONCLUSION

Our contributions for the study are two-fold. First, we described how a framework based on artificial intelligence techniques can be harnessed to inform the design of analytics and text



processing systems for generating personalized advertising copies. We developed SGS-PAC, adapted from techniques such as web element weight evaluation model, advertising words sentiment analysis, advertising copy template automatically generation, and personalized advertising copy generation based on template matching.

Second, we conducted an experiment to ascertain the validity of our proposed SGS-PAC. Analytical result suggests that consumers have a preference for personalized advertising copies generated by the SGS-PAC over generic ones that appear on retail platforms. With respect to recent design science theories, our research contribution represents an “improvement”: a novel and holistic solution to an enduring and yet elusive problem (Gregor and Hevner 2013).

Text analytics and user profiling technologies that support the automatic generation of personalized copies constitute an increasingly critical endeavor as comprehension is laid as the foundation for the use of artificial intelligence (AI) in advertising. The results of our work have important implications for various aspects of advertising. We present three examples here:

Advertising copies in the current environment are repetitive and featureless, which fail to assist consumers in quickly discovering the target product. Personalized advertising copy will enhance the efficiency of information search and help consumers better understand the characteristics of various products.

The creation of advertising copy is also faced with the problem of repeated labor. In the e-commerce environment, even for the same product, the demand for copywriting varies in different settings such as advertisement promotion, shop decoration, and activity registration. With SGS-PAC, copywriters and advertisers are relieved from this repetitive work. They will have more freedom and thinking time to release their creativity.

SAS-PAC can also help to save cost and create better advertising copies. Advertising copy is advancing with time and SAS-PAC matches the current trend. It can absorb massive high-quality content from talents and merchants and then transform it into excellent copywriting output through machine learning, enabling the formation of a virtuous cycle.

Like all research on text analytics, this study has limitations. First, even though our proposed SGS-PAC is rather generic, the experiment was conducted on Chinese advertising copies. For this reason, it might be interesting to compare the performance of the framework across different languages. Second, because it is difficult to source for consumers who fit into all four profiles, subjects were expected to play various roles during the experiment. A field experiment is hence necessary as the next logical step to replicate and/or refine the findings from this study on actual consumers that fit into the four profiles. Third, because our experiment was not conducted on actual ad exchange and retail platforms, we were not able to receive consumers' feedback and expand on the parameters for optimization. Furthermore, even though this research qualitatively illustrates the usefulness of SGS-PAC with personalized advertising copies, it will be interesting to examine more fine-grained advertising evaluation in the future.

## **REFERENCES**

Abbasi, Ahmed, Yilu Zhou, Shasha Deng, and Pengzhu Zhang (2018), "Text Analytics to Support Sense-Making in Social Media: A Language-Action Perspective," *MIS Quarterly*, 42 (2), 427-464.

- Ansamma, John, P. S. Premjith, and M. Wilsy (2017), "Extractive multi-document summarization using population-based multicriteria optimization," *Expert Systems with Applications*, 86, 385-397.
- Baker, William E., and Richard J. Lutz (2000), "An Empirical Test of an Updated Relevance-Accessibility Model of Advertising Effectiveness," *Journal of Advertising*, 29 (1), 1-14.
- Bargh, John A (2006), "What have we been priming all these years? On the development, mechanisms, and ecology of nonconscious social behavior," *European journal of social psychology*, 36 (2), 147-168.
- Becker, Ingo F., Marc Linzmajer, and Florian von Wangenheim (2017), "Cross-Industrial User Channel Preferences on the Path to Online Purchase: Homogeneous, Heterogeneous, or Mixed?," *Journal of Advertising*, 46 (2), 248-268.
- Bollen, Johan, Huina Mao, and Alberto Pepe (2011), "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena," *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pp. 450-453.
- Campbell, Colin, Leyland F. Pitt, Michael Parent, and Pierre R. Berthon (2011), "Understanding Consumer Conversations Around Ads in a Web 2.0 World," *Journal of Advertising*, 40 (1), 87-102.
- Cristianini, Nello, and John Shawe-Taylor. 2000. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.
- Delbaere, Marjorie, Edward F. McQuarrie, and Barbara J. Phillips (2011), "Personification in Advertising," *Journal of Advertising*, 40 (1), 121-130.

- Dief, Nada A., Ali E. Al-Desouky, Amr Aly Eldin, and Asmaa M. El-Said (2017), "An Adaptive Semantic Descriptive Model for Multi-Document Representation to Enhance Generic Summarization," *International Journal of Software Engineering & Knowledge Engineering*, 27 (1), 23-48.
- Gregor, Shirley, and Alan R. Hevner (2013), "Positioning and presenting design science research for maximum impact," *MIS Quarterly*, 37 (2), 337-356.
- Guariguata, L., D. R. Whiting, I. Hambleton, J. Beagley, U. Linnenkamp, and J. E. Shaw (2014), "Global estimates of diabetes prevalence for 2013 and projections for 2035," *Diabetes Research and Clinical Practice*, 103 (2), 137-149.
- Hassan, Ahmed, and Dragomir Radev (2010), "Identifying text polarity using random walks," *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics: Association for Computational Linguistics*, pp. 395-403.
- Hecht-Nielsen, Robert. 1992. *Theory of the backpropagation neural network*. In *Neural networks for perception*: Elsevier.
- Jiu-le, Tian, and Zhao Wei (2010), "Words Similarity Algorithm Based on Tongyici Cilin in Semantic Web Adaptive Learning System," *Journal of Jilin University (Information Science Edition)*, 28 (6), 602-608.
- Kietzmann, Jan, Jeannette Paschen, and Emily Treen (2018), "Artificial Intelligence in Advertising," *Journal of Advertising Research*, 58 (3), 263-267.
- Kumar, V., and Shaphali Gupta (2016), "Conceptualizing the Evolution and Future of Advertising," *Journal of Advertising*, 45 (3), 302-317.

- Liu-Thompkins, Y, and E.C. Malthouse (2017), "A Primer on Using Behavioral Data for Testing Theories in Advertising Research " *Journal of Advertising*, 46 (1), 213-225.
- Liu, Xia, Alvin C. Burns, and Yingjian Hou (2017), "An Investigation of Brand-Related User-Generated Content on Twitter," *Journal of Advertising*, 46 (2), 236-247.
- Malthouse, Edward C., and Hairong Li (2017), "Opportunities for and Pitfalls of Using Big Data in Advertising Research," *Journal of Advertising*, 46 (2), 227-235.
- McNair, Douglas M, Leo F Droppleman, and Maurice Lorr. 1992. Edits manual for the profile of mood states: POMS. San Diego: Educational & Industrial Testing Service.
- Parkinson, Brian, Peter Totterdell, Rob B Briner, and SA Reynolds. 1996. *Changing moods: The psychology of mood and mood regulation*. Longman.
- Rautray, Rasmita, and Rakesh Chandra Balabantaray (2018), "An evolutionary framework for multi document summarization using Cuckoo search approach: MDSCSA," *Applied Computing & Informatics*, 14 (2), 134-144.
- Saaty, Thomas L. (2008), "Decision Making with the Analytic Hierarchy Process," *International Journal of Services Sciences*, 1 (1), 83-98.
- Schultz, Don (2016), "The Future of Advertising or Whatever We're Going to Call It," *Journal of Advertising*, 45 (3), 276-285.
- Simsek, Atakan, and Pinar Karagoz (2018), "Wikipedia enriched advertisement recommendation for microblogs by using sentiment enhanced user profiles," *Journal of Intelligent Information Systems*, <https://doi.org/10.1007/s10844-018-0540-5>.
- Turpin, Andrew, and Falk Scholer (2006), "User performance versus precision measures for simple search tasks," *Proceedings of the 29th annual international ACM SIGIR conference*

on Research and development in information retrieval, Seattle, Washington, USA: ACM, pp. 11-18.

Velasquez, Mark, and Patrick T Hester (2013), "An analysis of multi-criteria decision making methods," *International Journal of Operations Research*, 10 (2), 56-66.

Zenil, Hector, Fernando Soler-Toscano, Kamaludin Dingle, and Ard A. Louis (2014), "Correlation of automorphism group size and topological properties with program-size complexity evaluations of graphs and complex networks," *Physica A Statistical Mechanics & Its Applications*, 404 (16), 341-358.