

A Study of Some Factors Affecting e-Advertising Development

Jantima Polpinij¹ and Ratanachote Thienmongkol²

Department of Computer Science¹ and Department of New Media²
Faculty of Informatics, Mahasarakham University
Mahasarakham 44150 THAILAND

{Jantima.p¹, Ratanachote.t²}@msu.ac.th

Abstract

Advertising has an influence on the success of business. At present, advertising has been reshaped from 'old media' to 'new media'. One way of reshaping based on new media is to transform advertising to be e-advertising. Although there are the modern tools to create a perfect advertisement, it may fail if customers' needs are not taken into consideration. Therefore, the major emphasis in this article is primarily empirical study of effects in e-advertising. Then, we employ data mining methodology to study the factors which attract customers. Using the well-known the Apriori algorithm, these factors turn out to be designing on the Web, the Web usability, information of product and service, and time consuming of search and use.

Keywords: e-advertising, data mining, the Apriori algorithm, new media

1. Introduction

From the past until now, advertising has had influenced on business success. In addition, advertising could be able to communicate with the customers in several channels. Then it can provide innovative marketing solutions for the modern marketplace. This is because the advertising directly affects customer behaviors in the case of purchasing products and services [1]. According to a famous company, the Coca-Cola Company, it continually spends money on advertising to support recognition of their products. In 1993, Coca-Cola spent more than \$150 million to keep its name in the forefront of the public's eye [2].

However, the advertising also has limits in communication to customers because there are many conditions based on redoubtable competition. Therefore, if the companies lack enough understanding to choose the appropriate way and tool for their target, it may fail [1]. It means, the message will not be communicated to their customers and it also creates investment cost. In the meantime, advances in information technology and communication have been reshaping the

advertising from 'old media' to 'new media'. Then "Old media is understood to mean the mass-media business models typified by television, radio or newspaper. Modern advertising has been strongly geared towards customers with several tools that are called "new media". In general, new media [4] refers to new forms of human and media communication that have been transformed by the creative use of technology to fulfill the same basic social need to interact and transact. At present, advertising has been moving to the Internet. It is a new way to directly and any more easily present information to the customers and improve organizational effectiveness [3].

Therefore, the state of the art of new media has been more affected with the advertising industry. Based on this, being e-advertising is becoming key to companies' success. This is because it can reduce funding costs in the area of advertising. Finally, interactive narratives are increasingly technologically possible and expected to affect e-business forms.

Actual implementations are not rare, but glitzy design has also been become significant. Specializing in custom web design, creative strategies, and integrated resources with experience are important to build the e-business and deliver effective results. Therefore, the major emphasis in this article is primarily empirical study of effects in e-advertising. Then we applied data mining methodology [5] to find the main factors of customers that are allured e-advertising. Our previous research proposed this method for automotive market segmentation analysis [5].

The rest of this paper is organized as follows. In Section 2, it is literature review. We describe the research methodology based on data mining in Section 3. Subsequently, some experimental results will be proved. The conclusion and future work are provided in the final section.

2. Literature Reviews

Due to competitive market, many companies' operators are constantly searching for alternative methods to supplement their income. The market is changed to appear on the Internet. In addition, the Internet becomes the new way to promote the products and services of each enterprise [3, 6].

Previously, data mining techniques have been used for intelligent e-business analysis. In 2001, Fayyad et al [7] addressed that the successful deployment of analytical solutions to e-business enterprise data requires data warehouse construction, efficient updates over multiple data sources, integration of data mining technologies with the underlying warehouse, and delivery of results in a form consumable by business end-users.

Based on above, advertising has also been changed the way of inducement to the Internet. It is so called 'e-advertising'. With our study, e-advertising spending has grown by 41%, making it now a larger part of the UK's advertising market than national newspapers [8]. Therefore, with marketing and advertising online enjoying meteoric growth, marketing executives at companies of all sizes are waking up to the fact they need to take the channel seriously [8].

However, established advertisement on the Internet should be carefully considered. Nielsen [9] gave a study which showed that only the top 0.01% of websites can generate sufficient revenues from advertising. In the larger picture, advertising is almost irrelevant for the success of the Web. Right now, e-advertising is attracting much media attention for two reasons: it is only the alternative way to generate a direct revenue stream, and it is hard to specify that this way is better than the old perspectives such as television, radio, or newspaper [9]. This is because although there are the modern tools to create a perfect advertisement, it may fail if customers' needs are not taken into consideration [10]. For example, what are the customers' needs from that e-advertising? And what is the characteristic of e-advertising that attracts the customers? These questions lead to watchfully reconsider about design and the modern tools in the new media age. That means the significance of e-advertising creation is not only the modern tools, but also the creative idea and design of e-advertising on the Web.

3. Research Methodology

This research proposes data mining to analyze some of the factors that attract the customers to purchase products and services through watching e-advertising. Then we collect data of using e-advertising from questionnaires. That means this research does not use any standard dataset. In addition, the Apriori algorithm, which is a data mining algorithm, is applied for factors analysis based on data mining method. Then, the research methodology can be presented in Figure 1.

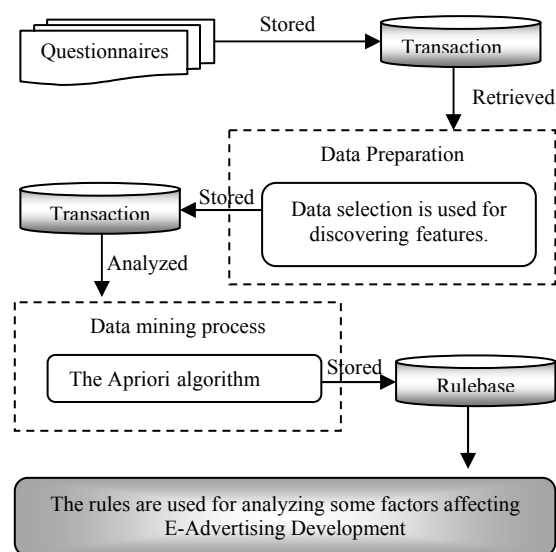


Figure 1: Research Methodology

3.1 Dataset

In this research, we do not use any standard data set but we used 200 questionnaires for our work. Then we can gather some of factors that attract the customers by e-advertising. These factors are presented in Table 1.

Table 1. The interesting factors of customers who are attracted by e-advertising.

Factor ID	Factor description
1	Using Tools for Web development (e.g. flash)
2	Designing on the Web (e.g. layout, color, images)
3	Usability (e.g. functions, easy to use)
4	Product and Service information
5	Time consuming of Search and Use
6	Ordering pattern

3.2 Data preparation

Data mining additionally involves data pre-processing, and results delivery. Data pre-processing [11] includes loading and integrating data from various data sources, normalizing and cleansing data, and carrying out exploratory data analysis. It is an often neglected but extremely important step in the data mining process. Preparing data for data mining has many ways to do, e.g. data selection, data cleaning, and data transformation.

This research used data selection for preparing data. Then data selection is an essential step in successful data mining applications. It is to select feature that has practical significance in many areas such as machine learning, and data mining. The objectives of this process include: building simpler and more comprehensible models, improving data mining performance, and helping to prepare, clean, and understand data. Then, this research also used for feature selection. In the final, all of features in Table 1 are selected for our research.

3.3 The Method by the Apriori Algorithm

The Apriori algorithm [11] is a technique in association rule mining. The major steps in association rule mining are: (1) Frequent Itemset generation and (2) Rules derivation. An itemset is a collection of one or more items. Then, a rule can be generated by the following:

$$\text{Association rule: } X \rightarrow Y \quad (1)$$

The Association Rules technique needs two parameters, minimum Support Threshold (minSupp) and minimum Confident Threshold (minConf). The minSupp is used for generating frequent itemsets and minConf is used for rule derivation. Then minSupp and minConf can be expressed as follows:

$$\text{Support } (X \rightarrow Y) = P(X \cap Y) \quad (2)$$

$$P(X \cap Y) = \frac{\#(\text{customers who are attracted to } X \text{ and } Y)}{\#\{\text{customer}\}} \quad (3)$$

$$\text{Confidence } (X \rightarrow Y) = P(Y | X) \quad (4)$$

$$P(Y | X) = \frac{\#(\text{customers who are attracted to } X \text{ and } Y)}{\#\{\text{customer who are attracted to } X\}} \quad (5)$$

In the case of the Apriori algorithm, it uses the downward closure property, to prune unnecessary branches for further consideration. It finds the

frequent set L in transaction D . The Apriori algorithm [11] can be expressed in the term of pseudo code follows:

(C_k : Candidate itemset of size k)

(L_k : Frequent itemset of size k)

- $L_1 = \{\text{frequent items}\}$
- For ($k = 1; L_k \neq 0; k++$) do
 - $C_{k+1} =$ candidate generated from L_k ;
 - For each transaction t in database do increment the count of all candidates in C_{k+1} that are contained in t
 - $L_{k+1} =$ candidates in C_{k+1} with minSupp
- Return $\cup_k L_k$;

In our work, based on Table 1, factor ID is itemset and frequent itemset is an itemset whose support is greater than or equal to a minSupp threshold. Then an example of our transaction and rules extraction can be shown as follows:

Table 2. An Example of our transaction

Transaction ID	List of factors that is chosen
T001	2, 4, 3
T002	2, 3, 4
T003	3, 2, 4
T004	3, 4
T005	2,1,4

Let minSupp be 2. Based on Table 2, the example of e-advertising development suggested by the Apriori algorithm can be expressed below:

Step 1: We find that each item forms a frequent 1 – items and the 1-candidate are shown in Table 3.

Table 3. The 1-Candidates

Item	Sup.	Remark
1	1	It is less than minSupp, so it is deleted.
2	4	-
3	4	-
4	5	-

Therefore, it can represent the frequent 1-item follows:

Table 4. Frequent 1-item

Item set	Sup.
2	4
3	4
4	5

Step 2: For each candidate set, we count their occurrences. By self joining itemsets, we get:

Table 5. The 2-Candidates

Item	Sup.	Confidence (followed by equation 5)
2,3	3	$(3/4) = 0.75$ (equal to 75%)
2,4	4	$(4/4) = 1.00$ (equal to 100%)
3,4	4	$(4/4) = 1.00$ (equal to 100%)

Finally, we get the 3-candidate that is presented in Table 6.

Table 6. The 3-Candidates

Item	Sup.	Confidence (followed by equation 5)
2,3,4	3	$(3/4) = 0.75$ (equal to 75%)

In the example above, it presents an example of suggestions that customers may be strongly attracted by information of product and service on the Web more than other factors (see Table 1.). If customers are firstly attracted by designing on the Web, they are also attracted by the Web usability at 75% of confidence. On the one hand, they can be also strongly attracted by information of product and service at 100% of confidence. In addition, if customers are firstly attracted by the Web usability, they are also strongly attracted by information of product and service at 100% of confidence. Finally, they can be attracted by designing on the Web, the Web usability, and information of product and service, respectively, at 75% of confidence.

4. The Experimental Results

After extracting factors that can be used for e-advertising development suggestions, the results are presented in Table 7.

Table 7. The experiment results at minSupp = 2

Item sets	Confidence (%)
2	100
3	75
4	100
5	50
2,3	75
2,4	100

Table 7. The experiment results at minSupp = 2 (cont'd)

Item sets	Confidence (%)
3,4	75
2,3,4	100
2,4,5	75
3,4,5	75
2,3,4,5	75

As above, it could be said that designing on the Web and information of product and service factors should be the primary components on e-advertising. That means the systems should not only suggest products and provide information to customers to help them decide which items to purchase, but also creative design on the Web. Based on Table 7, these factors show the confidence at 100%. In addition, these factors still are the main factors in each candidate. Then, the result coheres with a theorem of Cox and Dale [12]. They said that web design is the key to a successful web site, because perspective of web site can be like as an overall image of company. Also it can create a positive differentiation that customers can recognize and continue to access the web site. Then it depends on an appropriate combination of graphic design in terms of art such as color tactics, typographic design, multimedia technique, and so on. On the one hand, for information of product and service factor, they believed that the content refers not only to the category of products and services offering on the web site, but also the e-business policies. Therefore, it is also important to be contained on e-advertising [12].

Besides two factors above, the web usability is also important for e-advertising development. This is because it is possible that if the web is easy to use and it can support the customers' needs, e-advertising is interested by customers. Meanwhile, time consuming of search and use factor, if customers can access and look an e-advertising in very short time delays, they may not ignore the product and service on e-advertising. It can be referred to Gann's research [13], he found that the fast speed in download home page and other link has more influence with customer behaviors in term of keeping customers who always find any information through using the Internet. His report indicated that if a web page takes less than seven seconds to download information, fewer than 10% of customers will exit the web. If a web page takes eight seconds for access, 30% of customers will exit, while if it uses at least 12 seconds for access, 70% of customers will certainly exit the web. In final, it could be said that the common of web site

indeed are the convenience and speed in quality service [13].

5. Conclusion and Future work

The major emphasis in this article is primarily empirical study of effects in e-advertising. Then we applied data mining methodology to find the main factors of customers that are allured e-advertising. Then, using the well-known the Apriori algorithm, the main factors turn out to be designing on the Web, the Web usability, information of product and service, and time consuming of search and use. That means the significance of e-advertising creation is not only the information of product and service, but also the creative idea and design of e-advertising on the Web, access time, and usability of e-advertising. These factors are also significance for e-advertising development because these factors can help e-advertising's owner more understanding for their targets' need.

As above, it could be said that data mining method can be applied for factors affecting e-Advertising development. However, it is actually only an example in the case of data mining approach and the Apriori algorithms can be used for studying factors in this case. Then, these factors, extracted by our approach, are important components not only for e-Advertising, but also any e-commerce websites.

For our future work, we will apply data mining method and the Apriori algorithm to study important components of e-Advertising in the case of e-commerce website design, especially color specification for e-commerce. This is because color is the basic building material of two-dimensional images and visual experiences. The color analysis can play a pivotal role in the customer's critical decision [14] - to buy or not buy.

6. Acknowledgement

We are also grateful to thank Dr. Richard Booth, an expert in the area of knowledge representation of Faculty of Informatics, Mahasarakham University, who suggested for organizing this paper.

7. References

- [1] Clow, K.E. & Baack, D. (2004) *Integrated Advertising, Promotion, and Marketing Communications*, 2nd edn, Prentice Hall, New Jersey.
- [2] Kotler, P; Brown, L; Adam, S & Armstrong, G (2004) *Marketing*, 6th edn, Pearson Education, Australia.
- [3] Chisholm, R.F. (1988) *Introducing Advanced Information Technology into Public Organizations*, *Journal of Public Productivity Review*, 11(4), pp. 39-56.
- [4] __, *New Media* [Online], Wikipedia, http://en.wikipedia.org/wiki/New_media, Download on 22 August 2007.
- [5] Polpinij, J. (2004) *Automotive Market Segmentation by Machine Learning*, *The Proceeding on Artificial Intelligence and Applications (AIA)*, Innsbruck, Austria.
- [6] McCleary K.; Asubonteng P. & Munchus G. (1995) *The effects of advanced information technology on organizational design*, *Journal of Management in Medicine*, Emerald Group Publishing Limited, 12(4), pp 20-23.
- [7] Fayyad, U.; Rothleder, N. & Bradley, B.; (2001) *E-Business Enterprise Data Mining*, KDD Tutorial.
- [8] __, (2007) *Digital Strategy: Made to Measure*, *New Media Age*, pp. 21, London.
- [9] Nielsen, J. (1997) *Why Advertising Doesn't Work on the Web*, *Web Advertising*.
- [10] Gil, A. & Garcia, F. (2004) *E-Commerce Recommenders: Powerful Tools for E-business*, *ACM student Magazine*.
- [11] Han, J. & Kamber, M. (2001) *Data Mining: Concept and Techniques*, 2nd edn, Morgan Kaufmann Publishing, San Francisco, CA.
- [12] Cox, J. & Dale, B.G. (2002) *Key quality factors in Web site design and use: an examination*, *The Journal of International Journal of Quality & Reliability Management*, 19 (7), pp. 862-888 (Online Emerald).
- [13] Gann, R. (1999), *Every second counts*, *Competing*, pp.38-40.
- [14] __ (2007) *Color and Usability Matters* [Online], <http://www.colormatters.com/usability.html>, Download on 22 August 2007.