Adwords Keyword Set Selection Decision Support System Using AHP and TOPSIS Method

Sholikin Ady Chandra*, Edi Winarko2, Sigit Priyanta3
1Master Program of Computer Science; FMIPA UGM, Yogyakarta, Indonesia
2,3Department of Computer Science and Electronics, FMIPA UGM, Yogyakarta, Indonesia
e-mail: *sholikin.ady.e@mail.ugm.ac.id, 2ewinarko@ugm.ac.id, 3seagatejogja@ugm.ac.id

Abstract

CV. Gitani Creative Agency is a company engaged in the field of creative agency providing digital marketing service. Google Adwords is a platform used by the company to run this service. Keyword set selection is critical to the performance of ads. However, finding the right keyword set is not an easy task. The company needs to consider various criteria to get the optimal advertising results. Decision support system (DSS) is needed as an objective reference in the process of keyword set selection. The criteria for decision-making are click, impressions, cost, and avg CPC.

AHP method is used to compare the value of each criteria and then generate priority weights of each criteria. While TOPSIS method is used for alternative ranking. The combination of these methods aims to improve the performance of TOPSIS method.

The result of this study shows that the combination of AHP and TOPSIS methods can be used to determine the best keyword set for ads. Based on the testing results, DSS can do alternative ranking correctly in accordance with the results of manual calculation and it is also flexible to the changes in criteria and alternatives.

Keywords—decision support system, ahp, topsis, ranking, adwords
1. INTRODUCTION

The rapid growth in the number of internet users encourages companies to intensify their online advertising [1]. This type of advertising has shaped the internet world we know today. Displaying banners is a very prominent way to reach consumers, with more than five trillion banner ads being served each year [2].

Google Adword is a platform from Google providing advertising services. Using this platform, advertisers can create and manage advertisements to help promote their businesses, sell products or services, business branding and increase traffic to their websites [3]. The reach of the Google network is extensive, as it explores various market segments. Everyone knows Google has the most users as a website. Its average user every month can reach up to 271,534,144 people [4].

Ads created by advertisers will appear on Google’s search engine or network partners. Eligible ads to appear are based on similarities between Google users’ search keywords and the keywords from the Adword itself [3]. For representative results on Adwords, advertisements must contain a keyword set describing the ad unit [5]. Using a specific keyword set is recommended to obtain the right target market.

CV. Gitani Creative Agency is a company engaged in the field of creative agency, and one of the services provided is digital marketing. In serving the customers, this company utilizes the Google Adwords platform to advertise. Selecting a keyword set is a crucial part of the advertising process. The company has set five parameters or criteria for selecting a keyword set, including click, impressions, cost, and avg. CPC. The complexity of the alternatives and the criteria for determining the best keyword set might emerge as a distinct problem, especially if done manually.

Based on the aforementioned problems, we need a decision support system to support advertisers to get the best alternative of keyword set. The method used in the decision support system is the AHP method and the TOPSIS method. AHP method was selected as it is based on a comparison of two variables, allowing reduced inaccuracy risk in the weighting process [6]. According to Saaty, the AHP method can be used to solve complex problems, while considering many parameters or criteria [7].

The TOPSIS method is selected as its concept considers the best priority is not only seen from the shortest distance of a positive ideal solution, but also the longest distance of a negative ideal solution [8]. The TOPSIS method can help optimizing decision making process to solve practical decision problems, because the concept is simple and easy to understand. The computation is also efficient and has the ability to measure the performance of alternative decisions in a simple mathematical form [9]. In this study, the TOPSIS method is used for alternative ranking.

The combination of these two methods aims to improve the performance of the TOPSIS method. According to Shih et al, it requires initial weights (parameter weights) for further data processing [10]. In combining the AHP and TOPSIS methods, the weighting of parameters uses the AHP method through pairwise comparison and consistency check. Then the TOPSIS method is used for alternative ranking.

2. METHODS

This study aims to design and implement a decision support system for selecting a keyword set using the AHP and TOPSIS methods, to overcome the difficulties in determining a keyword set for advertising on Google Adwords. The design of decision support system is illustrated in Figure 1.
The first stage was analysing the criteria and alternatives for the keyword set needed for advertising. The next stage was using the AHP method to get the weight of each criterion. After that, the TOPSIS method was used to get the ranking value of several alternatives.

2.1 Determining the Criteria

Determination of the alternative order of the keyword set is based on several criteria. The criteria in this study were obtained based on interviews with the advertising team in the company. The interview sessions show that advertisers are very concerned about Cost and Impressions. In addition to the interview, the determination of criteria also considers the ones used in previous studies.

Arroyo-Cañada and Gil-Lafuente conducted a study aimed at ranking and selecting different alternative keyword sets to maximize the awareness and traffic to the website [11]. The alternative used is all the keyword sets that have been run for investment ads from stock brokers. This study used six criteria, including click, cookies week, OTS, CTR, Avg CPC and relevance.

After a review of the study by Arroyo-Cañada and Gil-Lafuente, the current researcher decided that the suitable criteria include click and Avg CPC. Cookies week and OTS cannot be used since that study used the data of the advertisements that were already running, while this current study used forecasting data. Forecasting data do not provide values for the Cookies Week and OTS criteria. It also determines the nature of the criteria. Criteria are beneficial if they provide benefits for decision makers, while they are cost in nature if they incur costs for decision makers. Based on the results of the interviews and literature study, it can be concluded that the criteria suitable for this current research include the ones presented in Table 1.
Table 1 Description of criteria

| No | Code | Criteria      | Criteria Type |
|----|------|---------------|---------------|
| 1  | C1   | Click         | Benefit       |
| 2  | C2   | Impressions   | Benefit       |
| 3  | C3   | Cost          | Cost          |
| 4  | C4   | Avg CPC       | Cost          |

The following are the explanations for each of the criteria in Table 1:

a. Click is the number of clicks received for an ad. Click shows how many times an ad is displayed on Google search pages and gets a click from web visitors. Clicks are included in the benefit category because the more number of clicks, the better the ad performance.

b. Impressions is the number of views for an ad. Impressions indicate the number of times an ad is displayed on a Google search page. Impressions are beneficial as the more the number of ad displays, the better the ad is running.

c. Cost is the total cost spent by advertisers for an ad. Ads are expected to have a high number of clicks and impressions with low costs so this criterion is a cost in nature. The lower the ad costs, the better the ad performance.

d. Avg CPC is the average cost-per-click charged to advertisers for ads that are run. Same with the cost criterion, Avg CPC is a cost in nature. The lower the Avg CPC, the better the ad performance.

2.2 Determining the Alternatives

At this stage, the alternatives are determined by decision makers. The alternative in this system is a keyword set that advertisers will later select for an ad on Adwords. In this study, an advertiser is a part of CV. Gitani Creative Agency. The company has various types of clients who want their goods or services advertised. Therefore, companies need a decision support system that is flexible against alternative keyword sets. Alternatives used in the decision support system can be changed depending on the nature of the projects.

Steps to determine alternatives taken by the CV. Gitani Creative Agency is as follows:

a. Discussing the goods or services to be advertised with clients.

b. Writing down the main keywords describing the ad.

c. Creating a keyword set derived from each of the main keywords obtained. Derived keywords are the ones that are more specific and still relevant to the main keywords. For example if the main keyword is "Muslim women's clothing" then the derived keywords can be "latest Muslim women's clothing, Muslim women's clothing prices, Muslim women's clothing in white" and so on. Derived keywords are obtained using the Semrush tool.

d. All the keyword sets created are used as alternatives for the decision support system.

For example, there is a client who want their Muslim women's clothing business to be advertised for 3 days starting from May 16, 2019 to May 18, 2019. The number of days is determined based on the amount of budget the client has. The budget allocated is Rp. 300,000 with a daily budget allocation of Rp. 100,000. Then, a research is conducted on various alternative keywords relevant to the purpose of advertising, in the steps as described above. The alternative keyword sets are presented in Table 2.
After the keyword set alternatives are obtained, the next step is to look for alternative performance data based on the selected criteria. Alternative performance data are obtained from forecasting data using the Keyword Planner for a span of 3 days starting from May 16, 2019 to May 18, 2019. The values of each alternative obtained from the Keyword Planner are presented in Table 3.

### Table 3 Alternative performance data

| Alternative Code | Click | Impressions | Cost    | avg CPC |
|------------------|-------|-------------|---------|---------|
| A1               | 76    | 2.150       | 68.438  | 900.5   |
| A2               | 85    | 2.353       | 80.772  | 950.26  |
| A3               | 60    | 1.714       | 54.211  | 903.52  |
| A4               | 80    | 2.221       | 71.062  | 888.28  |

### 2.3 AHP Method

Analytic Hierarchy Process (AHP) developed by Thomas L. Saaty is a method aimed at solving complex problems. What is complex is the parameters or aspects affecting quite a lot of problems. The complexity may be caused by unstructured problems, uncertainty of decision makers’ perceptions, or the unavailability of sufficient data. The steps to calculate the criteria weights include [12]:

1) Make a pairwise comparison matrix

The first step to deciding priorities is to make a pairwise comparison matrix, including elements compared in pairs against a specified criterion. If there is n criterion, a n x n matrix

---

*Adwords Keyword Set Selection Decision Support System Using AHP... (Sholikin Ady Chandra)*
will be obtained. To fill the matrix, we need a number illustrating the relative importance of an element to another element. Table 4 presents the rating scale for the pairwise comparison by Saaty, the scale defines and explains values from 1 to 9 set for consideration in comparing pairs of criteria.

| Intensity of importance | Definition                        |
|-------------------------|----------------------------------|
| 1                       | Equally important                |
| 3                       | Slightly more important          |
| 5                       | Much more important              |
| 7                       | Far more important               |
| 9                       | Extremely more important         |
| 2, 4, 6, 8              | The values between the two       |
|                         | considerations are close together|

2) Multiplying each element in the same line and the result is done by the square root of n according to equation (1)

\[
\tilde{w}_i = \sqrt{\prod_{j=1}^{n} a_{ij}}; \quad \text{dengan} \quad i = 1,2,\ldots, n
\]  

Information:
\( \tilde{w}_i \): weight of the i normalization criteria which has not been normalized
\( a_{ij} \): assessment of the importance of criterion i versus criterion j
\( n \): number of criteria

3) Normalize the square root to get the weight (eigen vector) according to equation (2)

\[
w_i = \frac{\tilde{w}_i}{\sum_{i=1}^{n} \tilde{w}_i}
\]  

Information:
\( w_i \): normalized criteria weights (eigen vector)
\( \tilde{w}_i \): criteria weights that have not been normalized

Stage 2 – Determining Logical Consistency

In decision making, it is important to know how good the consistency is, since we would not want the decision to be based on considerations with low consistency, thus seeming like random considerations. Finding consistency can be done by looking at the value of CR (Consistency Ratio). Following are the steps to calculate CR value [12]:

1) Calculating eigen value (\( \lambda_{maks} \)). \( \lambda_{maks} \) value is obtained using equation (3)

\[
\lambda_{maks} = \sum_{j=1}^{n} \left( \sum_{i=1}^{n} a_{ij} \right) w_j
\]  

Calculation steps:
a) Adding up the values per column from the pairwise comparison matrix.
b) Multiplying each value of the sum added from the column by the priority value corresponding to each criterion.
c) Adding up the results of the previous multiplies to obtain nilai \( \lambda_{maks} \) value.

2) Checking the Consistency Index using equation (4)
Consistency index: \( CI = \frac{\lambda_{\text{max}} - n}{n - 1} \) (4)

Information:
- \( CI \): Consistency index
- \( n \): The number of criteria
- \( \lambda_{\text{max}} \): The largest eigenvector value of an order matrix \( n \)

3) Calculating and checking the Consistency Ratio (CR) using equation (5)

\[ CR = \frac{CI}{RI} \] (5)

Information:
- \( CR \): Consistency Ratio
- \( CI \): Consistency Index
- \( RI \): Random Index

If the value is more than 0.1, the filling of pairwise matrix must be corrected. List of Random Indexes can be seen in Table 5.

| \( n \) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| \( RI \) | 0 | 0 | 0.58 | 0.9 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 | 1.49 |

2.4 Metode TOPSIS

TOPSIS is an MCDM method focusing on measuring the distance between the alternatives being proposed and two alternative of bipolar references [8]. The bipolar reference in question is a positive ideal solution and a negative ideal solution. TOPSIS requires decision making problems to be written in the form of a decision matrix. Each row represents an alternative and each column represents a criterion or evaluation parameter. Additionally, weight vectors must also be defined for each criterion.

According to Seçme et al, the TOPSIS principle states that the selected alternative must have the closest distance to a positive ideal solution (PIS) and the farthest from a negative ideal solution (NIS) from a geometric point of view using euclidean distance with optional weights of each attribute, in order to determine the relative proximity of alternatives to the optimal solution.

In general, the procedure of the TOPSIS method includes the following steps [9]:

a. Building a normalized decision matrix using equation (6)

\[ r_{ij} = \frac{x_{ij}}{\sum_{j=1}^{m} x_{ij}} \] ; with \( i=1,2,\ldots,m \); and \( j=1,2,\ldots,n \). (6)

b. Calculating a weighted normalized decision matrix using equation (7)

\[ y_{ij} = w_i r_{ij} \] ; with \( i=1,2,\ldots,m \); and \( j=1,2,\ldots,n \). (7)

c. Determining the positive ideal solution \( A^+ \) and the negative ideal solution \( A^- \) using equation (8)

\[ A^+ = (y_{11}^+, y_{21}^+, \ldots, y_{n1}^+) \]
\[ A^- = (y_{11}^-, y_{21}^-, \ldots, y_{n1}^-) \] (8)

Information:
- \( y_{ij}^+ \): Max \( y_{ij} \), if \( j \) is benefit attribute
- \( y_{ij}^- \): Min \( y_{ij} \), if \( j \) is cost attribute
\( y_{ij}^- \)
- Min \( y_{ij} \), if \( j \) is benefit attribute
- Max \( y_{ij} \), if \( j \) is cost attribute
A+ and A- to represent the most preferable alternative to the ideal solution and the least preferable in a row.

d. Calculating the distance of the value of each positive ideal solution and negative ideal solution to each alternative using equation 3.9 and equation 3.10

\[
D_i^+ = \sqrt{\sum_{j=1}^{n} (y_{ij}^+ - y_{ij})^2}; \text{with } i=1,2,...,m. \tag{3.9}
\]

\[
D_i^- = \sqrt{\sum_{j=1}^{n} (y_{ij}^- - y_{ij})^2}; \text{with } i=1,2,...,m. \tag{3.10}
\]

e. Determining the relative closeness of each alternative to each ideal solution using equation 3.11

\[
V_i = \frac{D_i^-}{D_i^+ + D_i^-}; \text{with } i=1,2,...,m. \tag{3.11}
\]

f. Ranking alternatives in a descending order

3. RESULTS AND DISCUSSION

This chapter covers the system testing and the discussion. The test for decision support system refers to its design explained in the research method section.

3.1 Testing the Results of Manual Calculations

The test results are manually calculated using Microsoft Excel to compare with the decision support system for set selection. This test aims to evaluate the results of system implementation in accordance with the system design described in the previous section.

3.1.1 AHP Calculation

AHP calculation test is done by comparing the results of systemic calculation to that of manual calculation. A comparison of the results of the calculation of criteria weights is presented in Table 6.

| Criteria         | Criteria weights                      | Manual calculation | Systemic calculation |
|------------------|---------------------------------------|--------------------|----------------------|
| Click            | 0.0861                                | 0.0861             |                      |
| Impressions      | 0.3072                                | 0.3072             |                      |
| Cost             | 0.1722                                | 0.1722             |                      |
| Avg CPC          | 0.4345                                | 0.4345             |                      |

As presented in Table 6, the results of the systemic calculation of criteria weights suggest the same value as that of manual calculation.

3.1.2 TOPSIS Calculation

TOPSIS calculation test is done by comparing the results of the systemic calculation to that of manual calculation. Comparison of calculation results is presented in Table 7.
Table 7 Comparison of the results of the calculation of TOPSIS method

| Alternative code | Relative Proximity Value | Manual calculation | Systemic calculation |
|------------------|--------------------------|--------------------|---------------------|
| A4               | 0.6490                   | 0.6490             |                     |
| A1               | 0.6164                   | 0.6164             |                     |
| A2               | 0.5721                   | 0.5721             |                     |
| A3               | 0.4181                   | 0.4181             |                     |

According to Table 7, the systemic calculation results have the same value as that of manual calculation.

3.2 Additional Testing

This test is necessary for the data used in the alternative assessment is forecasting data obtained from the Keyword Planner. It was done by comparing the results of systemic alternatives ranking of forecasting data to that of actual data. The actual data is obtained from the running ads on Google Adword with all the alternative keyword sets available for a certain period of time. All ad settings for each alternative keyword set are made the same. This test design is presented in Table 8.

Table 8 Additional Test Design

| Test design | Data period used          | Alternative keyword set |
|-------------|---------------------------|-------------------------|
| First case  | Three-day (16-18 may 2019)| A1, A2, A3, A4          |
| Second case | Two-day (27-28 may 2019)  | A1, A2, A3, A4          |
| Third case  | Two-day (27-28 may 2019)  | Hijab, Jilbab, Kerudung|

3.2.1 First Case

The forecasting data used is obtained from the Keyword Planner for three-day ad as used in the case study (May 16-18, 2019). While the actual data is obtained from three day-running ads on Google Adword with all the alternative keyword sets corresponding to the case study. Criteria and the value of the pairwise comparison use the same data as the case study's. Comparison of ranking results between forecasting data and actual data of running advertisements is presented in Table 9.

Table 9 First case test results

| Alternative Code | Forecasting Data | Actual Data |
|------------------|------------------|-------------|
|                  | Relative proximity value | Alternative ranking | Relative proximity value | Alternative ranking |
| A1               | 0.6164           | 2           | 0.6388 | 2          |
| A2               | 0.5721           | 3           | 0.4960 | 3          |
| A3               | 0.4181           | 4           | 0.4612 | 4          |
| A4               | 0.6490           | 1           | 0.7670 | 1          |

According to Table 9, the relative proximity has a different value for forecasting data do not have the exact same number as the actual data. However, the value is the same for alternative ranking of keyword sets, for both of forecasting data and actual data. This shows that forecasting data from Keyword Planner still has good accuracy.

3.2.2 Second Case

Alternatives, criteria, and pairwise comparison value between criteria use the same data as in the first additional test while the time period for advertising is different (two-day ad, 27-28 May 2019). A comparison of ranking results is presented in Table 10.
According to Table 10, the relative proximity has a different value for forecasting data do not have the exact same number as the actual data. However, the value is the same for alternative ranking of keyword sets, for both of forecasting data and actual data. This shows that forecasting data from Keyword Planner still has good accuracy.

3.2.3 Third Case

Criteria and pairwise comparison value between criteria use the same data as in the first and second additional tests, while keyword set alternatives used is different from the previous tests. The total alternative keyword sets used is three. To get the actual data of the ads run for two days (May 27-28 2019). Alternatives used in this test are presented in Table 11.

### Table 11 Alternative data testing third case

| No | Criteria Code | Alternative Keyword Set                                                                                                                                 |
|----|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1  | Hijab         | Hijab, Hijab cantik, Hijab terbaru, Hijab instan, Hijab segi empat, Hijab syari, Hijab pengantin, Hijab wisuda, Hijab kekinian, Hijab pesta         |
| 2  | Jilbab        | Jilbab, Jilbab instan, Jilbab syari, Jilbab organza, Jilbab rabbani, Jilbab segi empat, Jilbab pashmina, Jilbab saudia, Jilbab khimar, Jilbab rawis |
| 3  | Kerudung      | Kerudung, Kerudung rabbani, Kerudung terbaru, Kerudung instan, Kerudung syari, Kerudung segi empat, Kerudung organza, Kerudung saudia, Kerudung rawis, Kerudung khimar |

After obtaining the alternative forecasting and actual data, the research went on to ranking all keyword sets using AHP and TOPSIS methods. Comparison of alternative ranking results of actual and forecasting data in this test is presented in Table 12.

### Table 12 Third case test results

| Alternative Code | Forecasting Data | Actual Data |
|-----------------|------------------|-------------|
|                 | Relative proximity value | Alternative ranking | Relative proximity value | Alternative ranking |
| Hijab           | 0,5613           | 2           | 0,6212           | 2               |
| Jilbab          | 0,7148           | 1           | 0,7081           | 1               |
| Kerudung        | 0,2588           | 3           | 0,2633           | 3               |

This test shows that forecasting data from Keyword Planner still has good accuracy. According to Tabel 12 the relative proximity has a different value for forecasting data do not have the exact same number as the actual data. However, the value is the same for alternative ranking of keyword sets, for both of forecasting data and actual data.

With a decision support system, advertisers do not need to run all alternative keyword sets. Advertisers do not need to spend a large budget at a time. The allocated budget can be used to run a keyword set with the best performance, to enable the ads to run optimally. Advertisers simply need to run one of the best alternative keyword sets recommended by the system.
4. CONCLUSIONS

Decision support system built using a combination of AHP methods for weight calculation and TOPSIS for ranking the alternative keyword sets, is evidently able to produce alternative ranking recommendations. Additionally, based on additional testings, alternative keyword sets of forecasting data indicate the same ranking as the ranking result. Therefore, the decision support system can help advertisers optimize their budget for advertising using the best keyword set.

REFERENCES

[1] Interactive Advertising Bureau, “IAB Internet Advertising Revenue Report 2016,” *PriceWaterhouseCoopers, London*, 2017. [Online]. Available: https://www.iab.com/wp-content/uploads/2016/04/IAB_Internet_Advertising_Revenue_Report_FY_2016.pdf. [Accessed: 16-Jun-2018].

[2] A. Bleier and M. Eisenbeiss, “Personalized Online Advertising Effectiveness: The Interplay of What, When, and Where,” *Mark. Sci.*, vol. 34, no. 5, pp. 669–688, 2015.

[3] Google Support, “Google Ads: Definition,” *Google Ads: Definition*, 2018. [Online]. Available: https://support.google.com/google-ads/answer/6319?hl=en. [Accessed: 20-May-2018].

[4] Quantcast, “Top Websites,” *Top Websites*, 2018. [Online]. Available: https://www.quantcast.com/top-sites/. [Accessed: 20-May-2018].

[5] I. Conference, E. Technologies, I. Conference, and E. Technologies, *RECENT ADVANCES on EDUCATION and RECENT ADVANCES on EDUCATION and*, no. Eet. 2015.

[6] N. Komang, Y. Suartini, I. M. A. Wirawan, D. Gede, and H. Divayana, “DSS for " E-Private " Using a Combination of AHP and SAW Methods,” vol. 13, no. 3, pp. 251–262, 2019.

[7] A. A. Chamid and A. C. Murti, “Prioritization of Natural Dye Selection In Batik Tulis Using AHP And TOPSIS Approach,” vol. 12, no. 2, pp. 129–138, 2018.

[8] E. Roszkowska and D. Kacprzak, “The fuzzy saw and fuzzy TOPSIS procedures based on ordered fuzzy numbers,” *Inf. Sci. (Ny)*., 2016.

[9] S. Kusumadewi, *Fuzzy Multi-Attribute Decision Making (Fuzzy MADM)*. Yogyakarta: Penerbit Graha Ilmu, 2006.

[10] B. Singaravel and T. Selvaraj, “OPTIMIZATION OF MACHINING PARAMETERS IN TURNING OPERATION USING COMBINED TOPSIS AND AHP METHOD,” vol. 3651, pp. 1475–1480, 2015.

[11] F. J. Arroyo-Cañada and J. Gil-Lafuente, “A fuzzy asymmetric TOPSIS model for optimizing investment in online advertising campaigns,” *Oper. Res.*, pp. 1–16, 2017.
[12] T. L. Saaty, *Pengambilan Keputusan bagi para Pemimpin, Proses Hirarki Analitik untuk Pengambilan Keputusan dalam Situasi yang Kompleks*. Jakarta: Pustaka Binama Pressindo, 1993.