Tag Cloud Algorithm with the Inclusion of Personality Traits

Ahmad Affandi Supli  
College of Arts Sciences  
School of Multimedia  
University Utara Malaysia

Norshuhada Shiratuddin  
College of Arts Sciences  
School of Multimedia  
University Utara Malaysia

Azizi Ab Aziz  
Dept of Computer Science  
Faculty of Info. Technology  
University Utara Malaysia

ABSTRACT
It is imperative to consider human different perspective in order to visualize the information data towards users. Many studies proved that personality traits are one of the most significant factors that must be considered to give meaningful value when users see a view. This study tries to give ample evidence toward adjusting visual features on tag cloud visualization techniques. Since there is no study has tried to create an algorithm that can customize tag cloud visual properties based on personality traits. Therefore, the main objective of this study is to make tag cloud algorithm with the inclusion of personality traits by adjusting two prominent visual features (color and shape) as an integration of layout. In addition, the utilization of RBS (rule base system) approach as artificial intelligent method is also taken into account to make knowledge base that stores the relationship between the proper personality elements and particular layout. This paper also discusses findings from satisfaction evaluation of prototyping, which comprises three dimensions facet: overall layout, color, and shape. The findings showed that the majority mean value for each dimension is categorized in agree scale (6-point), which indicates that respondents are satisfied with the tag cloud layout display generated by proposed algorithm. The findings suggest interface designers to be careful in selecting the appropriate tag clouds layout to be displayed for users with varying personality differences.

General Terms
Algorithms, personality traits, information visualization

Keywords
Tag cloud, visualization, satisfaction, evaluation, forward chaining

1. INTRODUCTION
Tag cloud has emerged as one of the most famous visualization information techniques in recent years. Basically, tag cloud visualization technique interprets the usage of words within textual content, such as website, article, speeches or databases. In other word, the bigger words in tag cloud indicate that particular words are frequently mentioned in that website or article, whereas the smaller is on contrary. Tag cloud is also usually employed to portray frequency of tags that is created on users from social bookmarking websites [1] and also in social software [2]. Typically, the bigger word in tag cloud has escalation of color saturation in order to make it clearer.

The used of tag clouds have together appeared with many social software sites, such as Flickr website (photo sharing) and del.icio.us website (bookmark sharing) as well as Technorati website (blog searching). The tag cloud for these various sites is used to categorize information in the website itself (such as classifying photos or bookmarks). In addition, tags are mostly hyperlinks in which the user is guided to some certain websites that is signed on the tag. All these benefits have made tag cloud as a popular visualization technique, because it can help user to obtain high level of understanding towards data as well as exploration method [2].

In brief, in accordance with [2], there are four main functions that can be achieved by using tag cloud, namely searching, browsing, impression formation and presentation, and recognition or matching. Moreover, in recent years, many people start to use tag cloud visualization in recommendation system, because of the superiority of tag cloud that could show prominent terms towards users at a glance [3].

Nevertheless, to convey message through information visualization technique is not adequate without considering design guideline that requires comprehensive human perspective [4, 5]. It is so crucial to show visualization more convincingly upon users. Many studies have proved that personality type is one of human perspective factors to be considered in order to exhibit representation in meaningful ways towards viewers [6, 7, 8]. This is due to personality traits of users have significant influences that distinguish the people preferences one another. Therefore, some researchers suggested to customize the visual properties, such as layout, shape, structure, color, and so on, based on personality traits of each group of viewers.

A study has proved on interactive television prototype website that extroverted people personality types are much more associated with brighter color, on the other hand, the introvert are tend to choose sofer color or darker one [9]. Generally speaking, user who has some certain personality types more prefer to see a particular visualization layout compare to others [6, 8]. Therefore, the adjustment visualizations are now considered, because it will be able to improve the usage experience of users as well as their usability [10]. In addition, certain visualization display could also influence people feeling or emotion [10, 11].

Nevertheless, in general, most of tag cloud studies have not investigated yet the factors of personality types of users to yield customized tag cloud layout based on personality types. The earlier studies much more focused on examining different aspects of tag cloud usage. For instance, assessing tag cloud visualization by comparing to various interface design style [12, 13, 14], and the other is by inspecting visual features (position of tag, densities, size, and so forth) facet in tag cloud [15]. Moreover, a study that makes observations to compare different layouts of tag cloud and see its relationship towards
user’s performance for particular tasks. Lastly, in recent year, another study investigates performance on tag cloud tree style [16].

Despite tag cloud has many layout styles that are various across the internet, there is no algorithm yet that is intently created to manage customization of tag cloud properties based on personality types of user. Therefore, this study attempts to create a tag cloud algorithm, so that it will display correspondingly to users’ personality types.

To accomplish this, two main visual properties of tag cloud, which are color [17] and shape [18], are chosen. Particularly, the integration of these two visual features is customized to yield new tag cloud layout style based on personality types. Thus, the investigation on what kind of color and shape that are appropriately displayed based on personality traits are derived further. To begin, the discussion about personality traits is elaborated beforehand.

2. PERSONALITY TRAITS

Basically, in this study, it is significant to get insight about why people interact, prefer, behave differently towards something. This is due to the differences in their characteristic or personality. Personality can be defined as psychology mechanism system that is developed dynamically over the time and it has been affecting all patterns of behaviors, such as the way of taking decision, processing data and how they react on something [19].

To identify what kind of personality types people have, there are several instruments of questionnaires could be used, such as Big Five model [20], Myer Briggs Types Indicator (MBTI) [21] and so on. As described earlier, this study emphasizes on looking for the connection between two visual features (color and shape) and personality traits. The author found out that MBTI is an appropriate instrument to fulfill the objective of matching color and shape preferences across the literature.

Before coming out on discussion regarding what kind of certain color and shape are connected, it is vital to get more understanding about MBTI itself. The MBTI model was initially created by Katharine Cook Briggs and also her daughter (Isabel Myers Briggs) [21]. This model describes that people are grouped into 16 types in which each types has 4 scales. The scales are 1) Extroverted (E) or Introverted (I), 2) Thinking (T) or Feeling (F), 3) Sensing (S) or Intuition (N) and 4) Judging (J) or Perceiving (P). The detail descriptions towards those terms are elaborated below:

1. **Extroverted (E) versus Introverted (I):** This scale distinguishes people on how they behave in social daily live. Extroverted is assertive, sociable people and also is more outgoing. Meanwhile, Introverted is on the opposite way, which is more quiet and shy.

2. **Thinking (T) versus Feeling (F):** This scale explains on how people solve problem or take decision. If thinking type, they use their logical mind, whereby, feeling type tend to use their feeling or emotion involved in taking decision. This is due to they like to please people around them.

3. **Sensing (S) versus Intuitive (N):** This scale differentiates people on how they process given information. The sensing people are detail individuals which always utilize their all sense to process data. They also always like the things are being planned precisely on schedule and also they love routine activities. In contrary, the intuitive one doesn’t like things in so many details. They are big picture people, which mean they have capability to see patterns upon problem solving. They have subconscious intuitiveness to decide something.

4. **Judging (J) versus Perceiving (P):** The judging person means an individual who always manages their things in well prepared. This type is organized people. Whereas, the perceiving type is abstract one who is flexible towards their time. They are also spontaneous people.

These all elements in combination are summarized into 16 personality types. They are ISTJ, ISFJ, INFI, INTJ, ISTP, ISFP, INFP, INTP, ESTJ, ESFP, ENFP, ENTJ, ESFJ, ENFJ and ENTJ. In conjunction with this study, in fact the four elements as mentioned previously have significant linkages with study on color and shape. In other words, these four elements can be used to extract knowledge about both of color and shape to be displayed towards viewers based on their personality elements. Therefore, the next discussion explains more about those relations.

3. VISUAL FEATURES

3.1 Color Theory and MBTI

In this study, color is one of the visual features that are adjusted based on personality traits of users. Thus, the color theory called true color is selected to become part of knowledge base for tag cloud algorithm. This is due to this theory has direct significant relationship towards MBTI elements [22]. In other words, through extraction of MBTI personality element, the color preferences dominant can be detected. Initially, the developer of true color model was Don Lowry in 1978. This theory distinguishes people personality types into four primary colors (Orange, Gold, Green and Blue) (See Figure 1).

![True Colors](image)

**Fig 1: True Color**

The relationship between each other (True Color - MBTI) can be seen from the following figure (see figure 2). In figure 2, it is clearly defined that each color represents two similarities components which consist of each four personality types in MBTI. For instance, gold color is the representative of two combination element from Sensing (S) element and Judging (J) element. S and J itself are included in four personality types of MBTI. The other colors also represent each two different components. In the figure 1, it is also noticed that all correspond element of MBTI are varied from three scales MBTI, 1) Thinking (T) – Feeling (F), 2) Sensing (S) – Intuitive (N), 3) Judging (J) – Perceiving (P). Whereas, the scale Extroverted (E) – Introverted (I) are not involved in the relationship.
theory. The circle shape (NF – Intuitive and Feeling) is also equivalent with blue color MBTI elements and so on.

Fig 3: Psycho-Geometrics personality assessment tool

4. RULE BASE SYSTEM (RBS)

In order to incorporate those attributes which are gathered in aforementioned discussion (color and shape theory), a suitable artificial intelligence (AI) approach must be utilized. The Rule Base System (RBS) is selected to combine the attributes facts to generate the desired conclusion or decision. To achieve that, the RBS Forward chaining method is applied. Forward chaining is a sequence of chains that are being searched or passed over from a given problem in leading to gain its solution. Another definition is that forms of reasoning from facts aim to conclusion that is derived from existing facts [25].

This approach is appropriate to be applied as its distinctive feature “if-then” statement [25]. The “if” here means the condition is settled, whereby “then” statement means perform the action. In other words, all those facts actually already have appropriate solution that is comprehensively elaborated in this study. Each of evidence of gathering facts is accumulated as representative rule. This means if, for example, a particular condition of each rule is fulfilled, then the action given is fired. The fired rule has two possibilities of results, it may yield on firing another rule, or else it can result in the provided solution until all related rules are executed.

The RBS rules have two kind of inference, which are forward chaining and backward chaining [29]. Forward chaining is searching method that initiates the process of collection of data or facts, from the facts sought a conclusion that becomes a solution of the problems faced [29]. This method is also called data-driven due to it is started from data (facts) to reach goal (conclusion). Meanwhile, the backward chaining is on contrary, the searching process starts from the goal, the conclusion that a solution of the problems faced, to extracts the facts. In this study, the authors use forward chaining method, because the searching process starts from data to achieve appropriate conclusion (tag cloud layout based on personality traits).

The figure 4 depicts the forward chaining method that is formulated in this study. This forward chaining consists of bunch of rules about related knowledge in color and shape related to MBTI element. In figure 4, the concept of firing rules are applied from left side move forward to right side

Fig 2: Relationship between MBTI-True Color

However, many scholars agreed that Extroverted (E) – Introverted (I) scale has significant relation towards saturation of color. In accordance with [23] that reported from his experiments that Extroverted (E) in majority love brighter color, whereby Introverted (I) prefer lighter or softer color. In addition, a study from [9] also stated that extrovert people has preferences towards higher contrast color (saturated color), while on the opposite, the introvert people prefer lower contrast level of color (desaturated color). These all pieces of knowledge are then incorporated as knowledge base in this study to determine appropriate color based on personality elements that user has.

3.2 Shape Theory and MBTI

For another visual element (shape), the theory from Susan Dellinger called “Psycho-Geometrics” is applied. This is due to this approach also has conjunction with MBTI personality traits elements. Initially, Dellinger, who was fulfilling her Doctorate study in research of personality psychology and communication style, University of Colorado, succeeded developing this approach in form of 5 basic geometric shapes [24]. The shapes are circle, triangle, square, squiggle or zigzag and rectangle.

In this theory, it is explained that some people are more attracted to a particular shape due to personality type’s aspects [18]. In addition, the difference of brain usage also could determine the reason why people choose it. The left brain users prefer to choose linear shape form, such as triangle, square and rectangle. Meanwhile, the rest who prefer circle and zigzag shape are allocated as right brain user. The left brain users are creative thinker, using facts for processing data, and think in words. Whereas, the right brain users are creative thinker, using feeling and imagination to process data, and think in symbols or visualization.

The Figure 3 shows the relationships between MBTI personality element and shape form. As depicted, it is obvious that there is similarity between shape MBTI element and what the color MBTI element has. For example, the triangle shape has MBTI elements in Intuitive (N) and Thinking (T), and this is aligned with green color MBTI element types in true color

**True Colors Myers-Briggs Categories**

**Gold: SJ**

Includes: ISTJ, ESFJ, INFJ, and ESTJ

**Orange: SP**

Includes ESTP, ISFP, ESFP, and ISTP

**Blue: NF**

Includes: INFP, ENFP, INTJ, ENFJ

**Green: NT**

Includes: INTP, ENTP, INTJ

For another visual element (shape), the theory from Susan Dellinger called “Psycho-Geometrics” is applied. This is due to this approach also has conjunction with MBTI personality traits elements. Initially, Dellinger, who was fulfilling her Doctorate study in research of personality psychology and communication style, University of Colorado, succeeded developing this approach in form of 5 basic geometric shapes [24]. The shapes are circle, triangle, square, squiggle or zigzag and rectangle.

In this theory, it is explained that some people are more attracted to a particular shape due to personality type’s aspects [18]. In addition, the difference of brain usage also could determine the reason why people choose it. The left brain users prefer to choose linear shape form, such as triangle, square and rectangle. Meanwhile, the rest who prefer circle and zigzag shape are allocated as right brain user. The left brain users are creative thinker, using facts for processing data, and think in words. Whereas, the right brain users are creative thinker, using feeling and imagination to process data, and think in symbols or visualization.

The Figure 3 shows the relationships between MBTI personality element and shape form. As depicted, it is obvious that there is similarity between shape MBTI element and what the color MBTI element has. For example, the triangle shape has MBTI elements in Intuitive (N) and Thinking (T), and this is aligned with green color MBTI element types in true color...
leading to conclusion of color and shape. The all letter alphabets reflect all MBTI elements type.

![Fig 4: Forward Chaining of the rules](image)

To begin, first and foremost, the user is instructed to answer MBTI instrument test to extracting their score of all MBTI elements. Then, the comparison for all values are started to define his or her each element of each scale of MBTI personality elements.

Next part is to find the other fired rules to decide on which color that are appropriate based on their personality traits elements. The process continually goes forward until the other component, which are in terms of shape and saturated or desaturated color, are selected as it is triggered along the way.

5. PROPOSED ALGORITHM

As a result of the systematic process described earlier to yield appropriate combination of color and shape, the conversion upon forward chaining method is performed. The code below displays the algorithm in pseudo-code form.

Start

StartPersonalityTest

Calculate $E$

Calculate $I$

Calculate $S$

Calculate $N$

Calculate $T$

Calculate $F$

Calculate $J$

Calculate $P$

Get $E$

Get $I$

Get $S$

Get $N$

Get $T$

Get $F$

Get $J$

Get $P$

If $E > I$ Then

...
T = Thinking Element  
F = Feeling Element  
J = Judging Element  
P = Perceiving Element

**Personality1** = the basic scale of Extraverted (E) – Introverted (I)  
**Personality2** = the basic scale of Sensing (S) – Intuition (N)  
**Personality3** = the basic scale of Thinking (T) – Feeling (F)  
**Personality4** = the basic scale of Judging (J) – Perceiving (P)

This algorithm is now ready to be implemented in source code program. So that, to evaluate the satisfaction towards constructed tag cloud display from the proposed algorithm, case study must be conducted. In this study, the authors select career recommendation for youth as a case study in which the prototype will show the career recommendation to user in form of tag cloud display. Of course, to have a bunch of stored career recommendation, this study employs Multiple Intelligence (MI) theory as criteria that assess their prominent intelligences.

The founder of MI is a Professor of Education at Harvard University, named Howard Gardner [26]. He claims in his publishing book *Frames of Mind: The Theory of Multiple Intelligence* that human intelligences essentially compromises many point of view [27].

Initially, he proclaimed there are seven component of human intelligences, which are Verbal-linguistic, Logical-mathematical, Spatial-visual, Bodily-kinaesthetic, Musical, Interpersonal and Intrapersonal), and later on he added two more intelligences in his book, which is entitled “Intelligence Reframed” in 1999. There are Naturalistic and Existential, which are the additional intelligences that Gardner added, but the latter was not acknowledged by global acceptance [28].

This theory is implemented as career references in this study to come out with the recommended career results. Next, the tag cloud visualization is then applied to create such package of layout form that has particular color and shape based on proposed algorithm that is correspond to each user’s answer. The thing is that the prototype is equipped by two questionnaire test, first is MBTI test to extract user personality elements and second is Multiple Intelligent (MI) test to result his or her career.

However, as can be seen, the algorithm in this study is limited in terms of categorizing of color and shape to be integrated. The color of layout can be settled by change the font color, while in this study the layouts are produced by using absolute layout. The absolute layout means the layout is settled manually, this is due to the algorithm to create layout itself dynamically change is still insufficient across the internet.

6. INTERFACE DESIGN

The name of the prototype is YouthPDA. YouthPDA has login form in the first place, and then user can go further into system to fulfill provided questionnaire (MI and MBTI test). Both of MI and MBTI test have 2 options of language, which are Malay and English. After answering both questionnaire user is guided to see their personality types as well as their MI measurement scores. Eventually, by clicking process button, the user can see the recommendation result that has been covered by particular tag cloud layout. All possibility tag cloud layouts in the prototype that can be resulted by the algorithm can be seen in section 6.1 below.
To be noticed that the bigger font in tag cloud indicates the more prominent recommendation career result, whereby the smaller is not recommended but still useful as the function to make great impact of emphasizing upon the recommended one. In this study, as the definition of tag cloud, instead of the bigger words are generated by calculating tag that are available in textual content, the bigger words here is resulted from what the most prominent intelligent have been calculated from user. The font is then growing up to get the career result from database refer to particular intelligent variable.

7. SATISFACTION OF TAG CLOUD LAYOUT

The YouthPDA Prototype was tested to evaluate satisfaction of tag cloud display through prototyping method. The lab experiments were conducted to fulfill this purpose. There were 30 participants from University Utara Malaysia (UUM) instructed to fill up the questionnaire for data collection part. The questionnaire has three part of dimension of satisfaction. They are

i) **Overall Layout**: measured based on satisfaction of overall layout outlook (the combination of both shape and color) (7 items)

ii) **Color**: The users’ perspective of satisfaction in terms of color. (6 Items)

iii) **Shape**: The users’ perspective of satisfaction in terms of shape. (6 Items)

The evaluation of above-mentioned dimensions uses seven-point likert scales (see figure 14).

7.1 Findings and Analysis

There are 30 respondents involved in the experiment in which 12 people are male (40%) and 18 people are female (60%). In terms of education, undergraduate student are 11 people (36.7%) and postgraduate student are 19 people (63.3%).

By using descriptive analysis feature in SPSS software, the result for all dimensions measured in this evaluation is counted. Particularly, the mean score for each dimension is calculated in order to know the degree of satisfaction in average that respondents have.
The results for satisfaction evaluation of tag cloud display is shown in Table 1, whereas, the data visualization into graph is in figure 15. In the table 1, it can be seen that in general people are agree to have satisfaction toward tag cloud display by indication mean value is all above 5. In graph (see figure 15), all dimension are visualized more clearly.

Table 1. Overview of Mean Score for All Dimensions

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Layout</td>
<td>5.18</td>
<td>2.86</td>
<td>6.17</td>
</tr>
<tr>
<td>Color</td>
<td>5.51</td>
<td>3.17</td>
<td>6.83</td>
</tr>
<tr>
<td>Shape</td>
<td>5.42</td>
<td>2.67</td>
<td>7</td>
</tr>
</tbody>
</table>

As far as the satisfaction of the tag cloud displayed is concerned, dimension of overall layout, color and shape recorded the lowest score of 3 in average but with a small number of frequencies. In contrast, all three dimensions recorded the highest score of 6 and 7. As a result, this experiment shows a very impressive mean value of overall satisfaction. The mean value for each dimension that is greater than 5 indicated in general users are satisfied with tag cloud display that is generated by the proposed algorithm.

![Fig 15: Overview of Mean Score for All Dimensions in Graph](image)

8. CONCLUSION AND FUTURE WORK

Through systematic processes, this study has focused to construct an algorithm for tag cloud based on personality traits. Related theories have been reviewed as part of the proposed algorithm. The true color theory and psycho-geometry theory related to MBTI personality traits element, as well as rules base system approach are elaborated in the study. In particular, the forward chaining method is utilized in this research. Then, a case study to construct a prototype to evaluate the satisfaction towards tag cloud display was carried out. YouthPDA, the prototype, focuses on career recommendation result for youth. Lastly, the measurement of satisfaction based on result of three dimensions in a questionnaire (overall layout, color and shape) is analyzed by using descriptive analysis method.

The findings indicated that in general users are satisfied with tag cloud display that is generated by the proposed algorithm. It is therefore imperative to consider personality traits to customizing the related visual featured especially in tag cloud layout visualization. In the future, the authors will add other elements of visual properties of tag cloud visualization in order to get better results for satisfaction evaluation. Moreover, the algorithm will be constructed more dynamically by using database for storing the rules of knowledge.

9. REFERENCES


