A NEW ALGORITHM TO FIND MOST SIMILAR ITEMS WITH RESPECT TO TARGET ITEM BY USING AHP AND FUZZY MEMBERSHIP FUNCTIONS.

Murat ETÖZ
Karabük University, Department of Industrial Engineering, Karabük, Turkey
Erdinç SABUR
Karabük University, Department of Computer Engineering

ABSTRACT

Finding most similar commodities with respect to selected one is a core subject in today's competitive business environment. This paper introduces a new method and a new algorithm which can be used by commercial web sites or trading centres. The algorithm is based on finding and listing most similar goods/services related to customer preferences. In other words, when the user selects an item, this algorithm sorts other items depending on each item's similarity to the selected main item and lists them to the customer. The aims of the algorithm are to make customer selection easier, to increase his/her willingness to buy and to give a sense of satisfaction of sufficient searching before the buying decision. Our algorithm consists of 4 main steps. Determining item categories, finding item property set for each category, Finding priority weights of the properties by using Analytical Hierarchy Process, Defining fuzzy variables and calculating membership function values, using the values found on previous steps, calculating the similarity score of each alternative item with respect to the main item. An empirical study was done on printers category. Among the alternatives, the algorithm found the most similar item with %96,1 similarity score.

1. INTRODUCTION

Buying and selling goods, commodities via internet has been developed geometrically and reached billion dollars trading volume throughout the world during the past fifteen years. There are dozens of commercial web sites serving in many trading areas built for meeting customers' expectations and demands. Competition among commercial web sites necessitates several applications in order to attract attention and improve willingness of customer for buying goods or services. Advertisements on newspapers, billboards, televisions or highly visited web sites, competitive/aggressive price policies, listing most sold goods on the screen while customer looks at a specific product are some examples of these applications.

This paper introduces a new method and a new algorithm which can be used by commercial web sites or trading centres. The algorithm is based on finding

and listing most similar goods/services related to customer preferences. In other words, when the user selects an item, this algorithm sorts other items depending on each item's similarity to the selected main item and lists them to the customer. The aims of the algorithm are to make customer selection easier, to increase his/her willingness to buy and to give a sense of satisfaction of sufficient searching before the buying decision. A list containing items similar to a specific item helps the user to find closest alternatives. Generating the list of similar items is In case the user decides not to buy that item, he may decide to buy one of the items among that similar items list.

Determining if two items are similar or not, or more precisely, assessing the similarity of two items is a relative process in which the user's preferences must be considered. For example consider 2 cats. For a person who is not keen on cats, the two cats may seem quite similar. But a person who owns a pet cat and has observed cats for years can recognize many differences between two cats. Therefore expert knowledge had to be employed in similarity concept. The AHP, fuzzy membership function design and fuzzy AHP techniques are the tools that could be used in determining similarity between items.

The analytic hierarchy process (AHP) developed by Saaty [1,2] is widely used technique in modelling real life decision problems. AHP uses qualitative and quantitative information taken from experts and builds hierarchical framework among criteria by giving priority weights to each criterion. Alternatives are then weighted and sorted according to criteria. AHP also provides consistency checks for the pair wise comparisons.

AHP is a useful tool used in multiple criteria subjective decision making in which each criterion has different priority weights. AHP technique makes pairwise comparisons among criteria and uses eigenvectors to determine the weights of them.

If the set of ciriteria is represented as {1,2,…..,n} and A is a reciprocal matrix indicating preferences&judgements of evaluator among criteria (where $a_{i,j}= 1/a_{j,i}$ for $i\neq j$ and $a_{ii}=1$), then the eigenvector ω satisfies the equation $A\omega= \lambda_{max}\omega$ for $\lambda_{max}\geq n$. The elements normalised eigenvector ω represents the priority weights of criteria. The difference between n and λ_{max} represents the inconsistency of the judgements. If $\lambda_{max} = n$ then the matrix A is said to be consistent. In order to find degree of consistency, Saaty defined consistency index $CI=(\lambda_{max}-n)/(n-1)$. Saaty also calculated consistency indices of huge number of random matrices and derive consistency indices table for them. Saaty argue that if the ratio between consistency index on hand and consistency index from consistency indices table does not exceed 0,1, then the set of judgements which are represented in reciprocal matrix A is said to be consistent. If the ratio exceeds 0,1 then there is a inconsistency in evaluator judgements in doing pairwise comparisons and judgements have be revised.

Fuzzy set theory was first introduced by Zadeh [3]. Then Bellman and Zadeh [4] described the decision-making methods in fuzzy environments [4]. After that, an increasing number of studies were done about problems including fuzzyness by applying fuzzy set theory. [5-11]
The membership function was first determined by Zadeh [3]. It is the generalization of the indicator function in classical sets. Membership function is the function from a set to [0,1] interval. It represents the membership degree of an element x in fuzzy fuzzy set $\tilde{A}$. The value 0 means that x is not a member of the fuzzy set; the value 1 means that x is fully a member of the fuzzy set. The membership function is usually shown as $\mu_A(x)$ where $\tilde{A}$. is the fuzzy set, x is the element of fuzzy set and $\mu_A(x)$ is the dgree of membership of x to $\tilde{A}$. [3].

The rest of the paper consists of the following sections: In section 2, The similarity algorithm is introduced, In section 3, empirical study of the algorithm is presented. Section 4 contains conclusions and further studies.

2. Similarity Algorithm

Similarity algorithm consists of 5 main steps.
Part 0: Determining item categories (the set D). Finding item property set E_j for each category.
Part 1: Finding priority weights of the properties by using AHP.
Part 3: Defining fuzzy variables and calculating membership function values.
Part 4: Using the values found on above steps, calculating the similarity score of each alternative item with respect to the main item.

2.1. Abbreviations used in the algorithm

The abbreviations used in the algorithm explanation are as follows:
t: An item. An individual good that is suitable for commercial trade e.g. a BMW X6 SUV, a Xerox Phaser 3117 printer etc.
N: Total number of categories.
D: A set of categories of items.
G_j: the number of items in category j.
d_i: A category of items, e.g.:"TV sets", "printers", "washing machines" etc. i=1,..,G_j
E_j: The set of properties related with j'th category of items (d_j) where
S_j: Set of clusters related with j'th category of items (d_j) where
q_j: The number of clusters in S_j,($q_j = |S_j|$).
$e_{i,j}$: i'th element of E_j
$c_{j,k}$: A "cluster" that is a special set of properties where j=1,..,N and k=1,..,q_j.
For two integers a and b where 1≤a≤m and 1≤b≤m and $a \neq b$, $c_{j,k}$ is such a set that the following equations hold;

$$\bigcup_{k=1}^{m} c_{j,k} = E_j$$

and

$$c_{j,a} \cap c_{j,b} = \emptyset \text{ where } j = 1,\ldots,N \text{ and } 1 \le a \le q_j \text{ and } 1 \le b \le q_j$$

$P_{i,j,k}$: i'th property in the set $c_{j,k}$. An item property, can be such as "Speed", "Color", "height" etc.
$V_{i,j,k}$: The vector of all values of i'th property in k'th cluster of j'th category. This vector is a G_j dimensional vector.
$F_{i,j,k}$: The difference matrix of values of items in category j, wrt. thei'th property in k'th cluster. Both the rows and the columns of this matrix are the items.
$F_{i,j,k}[x,y]$: The element of the matrix $F_{i,j,k}$ at row x, column y.
$F_{i,j,k}[x,y]=0$ for x=y.
$r_{m,i,k,j}$= m'th item's i'th property value in k'th cluster in j'th category.
m_k: The total number of properties in k'th cluster, where k = 1..n.

2.2. SIMILARITY ALGORITHM

Part 0:

1) Determine the category D which will be studied on.
2) Determine the set of properties E of category D.
3) Determine the set of clusters S of category D.

Part 1:

1) Questionnaires containing pair wise comparison of clusters with respect to similarity criteria are conducted.
2) For each cluster, a questionnaire containing pair wise comparison questions of that cluster elements with respect to similarity criteria is conducted. Elements consist of properties and sub properties.

Part 2:

1) Using the questionnaire results and employing AHP technique, priority weights of all the elements are calculated.

Part 3:

This part is performed individually for each item in the category.

1) Partial similarity score for each property is found. In order to find the scores, quantization technique explained in section 2.3 is employed.

2) In each cluster, similarity score of each property in that cluster is multiplied by the weight of that property (found in part 2). If the property is multidimensional property and contains sub properties, then weight of each sub property is multiplied by similarity score of each sub property. The multiplication results are summed. This sum can be called as "the cluster similarity score".

3) Each cluster similarity score is multiplied by that cluster's weight (found in part 1). This result can be called as "the final similarity score" of a particular item.

4) A list is generated by sorting all items in the category other than the main item with respect to each item's final similarity score. The first item in this list being the most similar item and the last item being the least similar item to the main item.

Part 4:

2.3. Quantization:

In part 3.1 of the algorithm, In order to find a similarity score, a partial similarity score for each property is found. Partial similarity scores are calculated by using different calculation techniques. Depending on the necessary partial similarity score calculation technique, properties are classified into several types as below:

2.3.1. Types of Properties:

a) Simple valued property type: This type of properties have natural number Eg. "Miles per gallon" for "cars" category.
b) Type code number type: If the item has a property related with type, this property takes values in the range [1..N] (N being the number of all possible types of all the items in the category).
c) Simple Boolean property type: This is a property bearing a value showing if a feature is existing or not. eg. "Existence of short washing cycle" in "washing machines". These properties have either 1 (Existing) or 0 (Not existing) as a value. Eg. In "Printers" category, "Existence of a screen" property is an example.
d) Binary number property type: This type of property has a vector of Boolean values and a set of features. The Boolean values of this vector are represented as the digits of a binary number. Each digit holds a value either 1 or 0, representing either the existence or absence of a feature. While defining this type of property, first of all, the set of all possible features related with the property is found. Then, a binary number having 1 digit for each possible feature is defined. At the end, a binary number representing property features' existence is obtained. The number of digits of this binary number is equal to number of all possible features of the property in the considered category.

For example; printer media type is a binary number property. The features of this property are the set of all media types supported by at least one printer in the "Printers" category. This set includes "Plain paper"(1), "Ink jet paper"(2), "Photo paper"(3), "Envelopes"(4), "transparencies"(5), "Labels"(6), "Cards"(7), "Iron-on transfers"(8), "Borderless media"(9). There are other media types not included in this set. Those other media types are out of our interest because, any printer in the "Printers" category in our sample case, does not support them. We include the media types that are supported by at least one printer in the "Printers" category. For a particular printer that supports only plain paper and envelopes, the binary number representing Boolean vector will be 100100000.The first component of this Boolean vector, or in other words, the most significant digit of this binary number, denotes that this printer supports "Plain paper". Similarly, the fourth component of the Boolean vector or the fourth most significant digit of the binary number is 1 meaning that this particular printer supports "Envelopes". All other zeros mean that this particular printer does not support all other printer media types ("Ink jet paper"(2), "Photo paper"(3), "transparencies"(5), "Labels"(6), "Cards"(7), "Iron-on transfers"(8) and "Borderless media"(9)).

e) "Multidimensional property with binary number sub properties" type: Each sub-property is treated as a binary number type property; as the method suggests, a binary number will be defined for each sub property. We call this binary number as $\mathbf{B_{u,i,k,j}}$ for the u'th sub property of the i'th property in the k'th cluster in the j'th category. The number of digits of $\mathbf{B_{u,i,k,j}}$ is equal to the elements of the universal set which contains all possible values of the sub property. For example, assume that "memory card options" property (Let's say, this is the 3rd property in 2nd cluster at 5th category) has three sub properties which are named as "SD card" (1st sub property), "memory stick" (2nd sub property), "XD card" (3rd sub property). Each sub property has different capacity options. If a particular printer supports the 1 GB, 2 GB and 4 GB capacity SD cards, and if universal set of all possible card capacities in this printer category is represented as U={1GB, 2 GB, 4 GB, 8 GB, 16 GB, 32 GB, 64 GB}, then the binary number which represents "SD card" sub property for a particular printer is $\mathbf{B_{1,3,2,5}}$ = 1110000. The first 3 digits of the binary number represent the existence of 1 GB, 2 GB and 4 GB memories. The subsequent zeros represent the nonexistence of 8 GB, 16 GB 32 GB and 64 GB memories. The important point here is that binary number $\mathbf{B_{u,i,k,j}}$ exists if and only if the corresponding sub property exists. For example if a particular printer does not have a SD card sub property, $\mathbf{B_{1,i,k,j}}$ does not exist for all i, k, j.

f) Multidimensional property with simple valued sub properties: This is a binary number property type having a simple value for each of its sub-properties. Every digit of this binary number is associated with a sub-property. Existence or validity of a sub-property determines the value of its associated digit. If a sub-property is valid or existing in an item, the binary digit associated with that sub-property is 1 and 0 otherwise.

2.3.2. Designing Fuzzy Membership Functions of Properties

Fuzzy membership functions are designed in each category separately. The method of designing fuzzy membership functions for a category is explained below.

Since similarity to a specific item differs from similarity to another item, a different fuzzy membership function is designed for assessing similarity to each item with respect to each property.

For each property type explained above, a different design is devised as explained below:

a) Simple tangible valued property type fuzzy membership function design:

i) Property values of all items in a category are determined. For each property, the set of possible/alternative values is determined ($V_{i,j,k}$). For example, the thickness property is one of the properties in the selected category of items. The set of thickness values of all the items in that category

is determined such that if "thickness" property is the 3rd property of 7th cluster of the 4th category and if there are 7 items in the 4th category,

$V_{3,7,4}$= {1.2, 1.8, 2.3, 4.5, 6.7, 8.0, 11.21}

ii) The set found in step i is ordered from least to biggest.

iii) For each item in the category the differences of the item's value and all other items' values are found. This yields the difference matrix $F_{i,j,k}$ (i: property number, j : category number, k: cluster number).

iv) The maximum elements of each row of $F_{i,j,k}$ are found as:

$$\beta_x = \max_x \mathbf{F_{i,j,k}[x,y]}$$

v) If β ≠0, the fuzzy membership function is as follows:

$$\mu_x(\mathbf{r_{y,i,j,k}}) = 1 - \frac{|\mathbf{F_{i,j,k}[x,y]}|}{\beta_x}$$

vi) If β = 0, the fuzzy membership function is as follows:

$$\mu_x(\mathbf{r_{y,i,j,k}}) = 1$$

b) Type code number type fuzzy membership function design:

i) Property values of all items in a category are determined.

ii) For two integers x and y, such that 1 and

$$\mu_x(\mathbf{r_{y,i,j,k}}) = \begin{cases} 1 & if\ \mathbf{r_{x,i,j,k}} = \mathbf{r_{y,i,j,k}} \\ 0 & if\ \mathbf{r_{x,i,j,k}} \neq \mathbf{r_{y,i,j,k}} \end{cases}$$

The function designed above is the similarity fuzzy membership function for the x'th item of category j. This design is used for all items in the related category (j), with x having all integer values from 1 to G_j , thus the number of fuzzy membership functions designed for this property will be G_j.

c) Boolean valued property type fuzzy membership function design:

i) Property values of all items in a category are determined.

ii) For two integers x and y, such that 1 and

$$\mu_x(\mathbf{r_{y,i,j,k}}) = \begin{cases} 1 & if\ \mathbf{r_{x,i,j,k}} = \mathbf{r_{y,i,j,k}} \\ 0 & if\ \mathbf{r_{x,i,j,k}} \neq \mathbf{r_{y,i,j,k}} \end{cases}$$

The function designed above is the similarity fuzzy membership function for the x'th item of category j. This design is used for all items in the related category (j), with x having all integer values from 1 to G_j , thus the number of fuzzy membership functions designed for this property will be G_j.

d) Binary number property type fuzzy membership function design:

i) The boolean value vectors (binary numbers) of all the items are determined.

ii) For each item, the item's binary number is compared with all other items' binary numbers. While comparing two items' binary numbers, the number of the differing digits is counted.

vii) A matrix showing the number of differing digits is designed. This yields the difference matrix $F_{i,j,k}$ (i: property number, j : category number, k: cluster number).

viii) The maximum elements of each row of $F_{i,j,k}$ are found as:

$$\beta_x = \max_x \mathbf{F_{i,j,k}[x,y]}$$

ix) If β ≠0, the fuzzy membership function is as follows:

$$\mu_x(\mathbf{r_{y,i,j,k}}) = 1 - \frac{|\mathbf{F_{i,j,k}[x,y]}|}{\beta_x}$$

x) If β = 0, the fuzzy membership function is as follows:

$$\mu_x(\mathbf{r_{y,i,j,k}}) = 1$$

e) "Multidimensional property with binary number sub properties" type fuzzy membership function design: a fuzzy membership function for each sub-property is designed in a way the same as binary number property type fuzzy membership function design as mentioned above.

g) "Multidimensional property with simple valued sub properties" type fuzzy membership function design: The binary numbers of the main item and each alternative is compared. If both the main item and an alternative item have zeros in a binary digit (which refers to the same sub property), then the similarity score of that sub property is 1. If one has 0 and the other has 1 then the similarity score of that particular sub-property is 0. If both have 1s then the similarity score of that that sub-property is calculated in the same way as in a "simple valued property" type property.

3. AN IMPLEMENTATION:

The above algorithm is implemented on printers. For the printers category, we run our algorithm to find similar printers ranked from the most similar to least similar ones with respect to a particular printer selected. The selected printer is HP photosmart C4680 multifunction printer in our case. In order to simplify the demonstration of algorithm run, in our empirical study, we defined our printer database consists of 10 different printers. Eventually, in real life applications, the users of algorithm have to deal with much bigger databases. However, the working principle remains the same as in small sample case.

Algorithm:

Part 0:

1) D=Printers

2) We defined 36 different properties which determines the almost all of the specifications of the printers in our database. The property set E is as follows;

E={ Maximum Scan Resolution, Max. Document Size for Scanning, Two Side Scanning Ability, Maksimum Copy Resolution (Color), Maksimum Copy Resolution (Black), Print Quality (Color), Print Quality (Black), Print Technology, Print Speed (Black A4), Print Speed (Color A4), Copy Speed (Black A4), Copy Speed (Color A4), Memory Card Options, Memory, Power, Interface and Connectivity, Media feeder capacity, Media Types, Media Sizes, Two Side Printing option, Screen, Weight, Fax Ability, Color, Volume, Warranty Duration, Spare Part Support After Warranty, On Place Service, Call Center, Brand Name, Price, One Installment Amount, Third Party Cartridge refill options, Cartridge cost (black), Cartridge cost (color), Initial cartridge level percentage}

3) Properties defined above were then grouped into 7 clusters. The cluster set S is determined as;

S={Scan Properties, Print Properties, Speed Properties, Electronic Properties, Usage Conveniences, Service Conditions, Brand, Financial Properties}

The properties are assigned to these clusters as below:

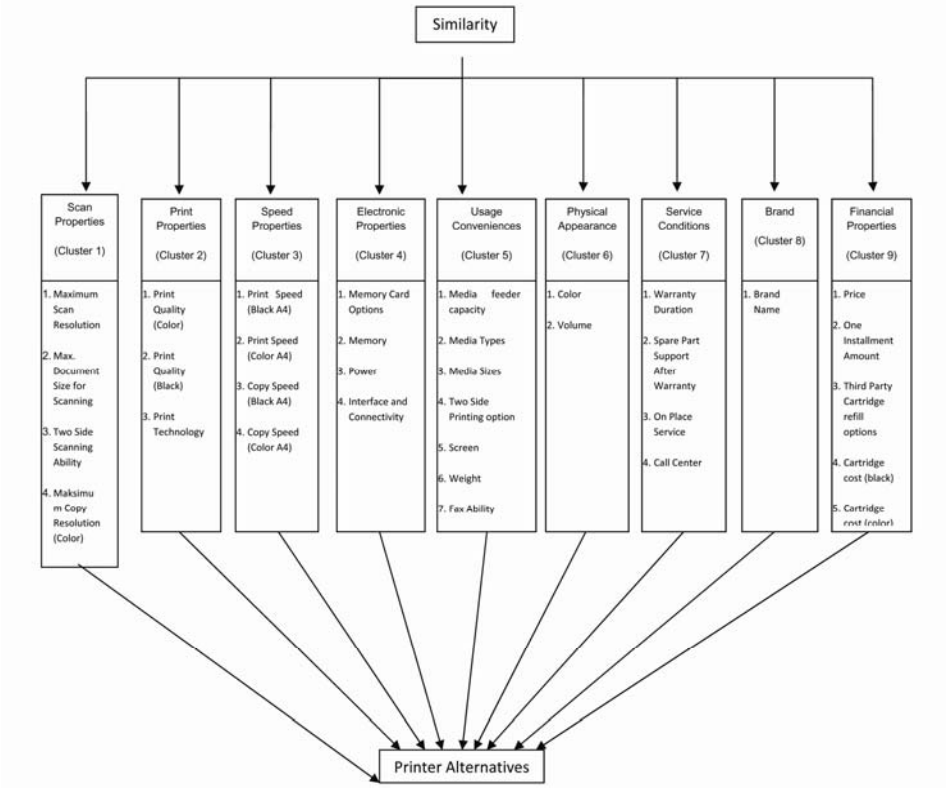
1. Scan Properties = { Maximum Scan Resolution, Max. Document Size for Scanning, Two Side Scanning Ability, Maksimum Copy Resolution (Color), Maksimum Copy Resolution (Black), }
2. Print Properties = { Print Quality (Color), Print Quality (Black), Print Technology }
3. Speed Properties = { Print Speed (Black A4), Print Speed (Color A4), Copy Speed (Black A4), Copy Speed (Color A4)}

a. Electronic Properties = { Memory Card Options, Memory, Power, Interface and Connectivity}

4. Usage Conveniences = { Media feeder capacity, Media Types, Media Sizes, Two Side Printing option, Screen, Weight, Fax Ability}
5. Physical Appearance={ Color, Volume }
6. Service Conditions = { Warranty Duration, Spare Part Support After Warranty, On Place Service, Call Center}
7. Brand = { Brand Name}
8. Financial Properties = { Price, One Installment Amount, Third Party Cartridge refill options, Cartridge cost (black), Cartridge cost (color), Initial cartridge level percentage }

The AHP network of the system is shown in Figure 1.

Figure 1. AHP network of similarity of printers



Part 1:

As part of the AHP method, questionnaires containing pairwise comparisons of clusters is filled out by 5 printer experts. These experts are selected among people having IT related jobs. Some of the questions asked in these questionnaires are as follows:

Is a pair of two printers more similar when their print speeds are similar (a) or when their print resolutions are similar (b)?

Which one is more important?

(a) is more important

(b) is more important

(c) both printer is equal

How much?

1) Extremely more important

1) Very strongly more important

5) Strongly more important
7) Moderately more important
9) Equal
Enter your grade here (1 – 9):

When the experts get into trouble deciding, the question is rephrased eg.:

In order to assume two printers are similar, is either similarity of their print resolutions (a) or similarity of their print speeds (b) more important.

Part 2:
At the end of AHP calculations, priority weights of each property is obtained as shown in Table 2;

Table 2: Weights of properties

Property	Weights
Brand name.	0,024198
Interface and connectivity	0,050505
Memory card options	0,029890
Power	0,003970
Memory	0,007272
Third party cartridge refill options	0,064328
Price (USD exld. Wat)	0,035451
Initial cartridge level percentage	0,005087
Cartridge cost black (USD excld. Wat)	0,013261
Cartridge cost color (USD excld. Wat)	0,012124
One installment amount.	0,035451
Colour.	0,002643
Volume (mm3)	0,013315
Print Quality (Colour)	0,036673
Print Quality (black)	0,037099
Print technology	0,112600
Two side scanning ability	0,010755
Maximum Scan Resolution	0,046396
Max. doc size for scanning(mm2)	0,090514
Maximum Copy resolution (black)	0,019352
Maximum Copy resolution (color)	0,019352
Spare part support after warranty	0,003801

On place service	0,006584
Call Center	0,002134
Warranty Duration.	0,011691
Print Speed (Black A4 ppm)	0,058243
Print speed (Colour A4 ppm)	0,050469
Copy Speed (Black A4)	0,015611
Copy Speed (Colour A4)	0,015611
Media Feeder Capacity	0,006721
Media Types	0,019717
Media Sizes	0,016652
Two side printing option	0,019717
Screen	0,014737
Fax ability.	0,081340
Weight.	0,006736

Part 3:
In this step the similarity score of each alternative printer with respect to the selected printer (HP photosmart C4680) is calculated. The property values of each printer are shown in Table 3. Differences of property values between HP C4680 and alternative printers are shown in Table-4. Partial Fuzzy membership function values of each printer with respect to properties($\mu_x(r_{yiijk})$) are shown in Table-5. The final similarity score are shown in Table-6.

Part 4:
The most similar printer to HP Photosmart C4680 is found to be HP Photosmart C4780 with 96,1% similarity score. It is followed by HP Deskjet F2420 (79,7%), HP Officejet J4500 (69,5%) and Canon Pixima MP490 (63,4%) respectively. The least similar printer is found to be Samsung ML1660 laser printer with %32,3 similarity score.

4. CONCLUSIONS

Similarity algorithm introduced in this paper puts a new approach of making comparisons between similar items. Algorithm provides degree of similarity between the items in quantitative terms. Quantitative results facilitate customer decisions of selecting most appropriate items among huge amount of alternative items. Basically algoritm makes all possible comparisons between items with respect to given criterias, weights them, sorts them, and reports the degree of similarity to the customers.

5. REFERENCES

[1] T.L. Saaty, The Analytic Hierarchy Process, McGraw-Hill, NY, 1980.

[2] T.L. Saaty, A scaling method for priorities in hierarchical structures, J. Math. Psychol. 15 (1977) 234–281.

[3] L.A. Zadeh, Fuzzy sets, Inform. and Control 8 0965) 338-356.

[4] R.E. Bellman, L.A. Zadeh, Decision-making in a fuzzy environment, Manage. Sci. 17 (1970) 141–164.

[5] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning. Part 1, Informat. Sci. 8 (1975) 199–249;

[6] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning. Part 2, Informat. Sci. 8 (1975) 301–357;

[7] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning. Part 3, Informat. Sci. 9 (1975) 43–80.

[8] S.M. Chen, A new approach to handling fuzzy decisionmaking problems, IEEE Trans. Systems, Man, Cybernetics 18 [1988)

[9] P.S. Das, Fuzzy groups and level subgroups, J. Math Anal Appl. 84 (1981) 264-269.

[10] G.J. Klir, T.A. Folger, Fuzzy Sets, Uncertainty, and Information, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1994.

[11] A. Rosenfeld, Fuzzy groups, J. Math Anal. Appl. 35 (1971) 512-517. 1012-1016.

Table-3: The descriptions of properties and property values of printers

Property name	Property Description	Cluster Name	Property Type	Property Type Description	HP C4680	Samsung ML1660	HP Photosmart C4780	Canon PIXMA MP250	Brother HL-2040	HP Deskjet F2420	6 HP laserjet ProP1102	Canon Pixma MP490	HP Officejet J4500	Samsung SCX4600
Brand name		Brand	Type code Number	Every brand has a numeric code HP: 1	1	2	1	3	4	1	1	3	1	2
Interface and connectivity	The versatility of types of printer-pc connections, e.g. usb cable, Firewire cable, parallel port cable	Electronic Properties	Binary number	5 digit binary number: 1st digit belongs to USB 2.0, second digit belongs to wireless, Third digit belongs to Firewire, 4th digit belongs to parallel port and last digit belongs to 10/100 ethernet	10000	10000	11000	10000	10010	10000	10000	10000	10001	10000
Memory card options	The versatility of types and sizes of memory card the printer can use, e.g. SD Card slots, memorystick card, USB slot Flashdisk, etc	Electronic Properties	Binary number	6 digit binary number: 1st digit belongs to memory stick, second digit belongs to memory stick duo, Third digit belongs to SD card, 4th digit belongs to SDHC, 5th digit belongs to all picture card and last digit belongs to USB slot for flash disks	111110	000000	111110	000000	000000	000000	00000	111111	000000	000000
Power	Maximum power requirement (watt) at operation mode	Electronic Properties	Simple Valued		20	250	20	11	450	22,8	360	12	9,6	686
Memory	The size of the printer's memory, eg. 32 MB or ram	Electronic Properties	Simple Valued		64	8	64	2	8	2	2	2	64	64
Third party cartridge refill options	whether the cartridge can be refilled by third party refillers or not	Financial Properties	Simple Boolean		1	0	1	1	1	1	1	1	1	0
Price (USD exld. Wat)	Purchasing price of the printer	Financial Properties	Simple Valued		68,5	88,5	72	46,5	87	59	111,5	70,5	119,9	164,9
Initial cartridge level percentage	How full is the cartridge provided included in the printers price	Financial Properties	Simple Valued		0,5	0,5	0,5	1	1	1	0,5	1	0,5	0,5
Cartridge cost black (USD excld. Wat)	The price of a new cartridge in the printer's own brand	Financial Properties	Simple Valued		18,49	70,99	18,49	15,99	58,99	18,49	79,5	15,99	18,49	89,99
Cartridge cost color (USD excld. Wat)	The price of a new cartridge in the printer's own brand	Financial Properties	Simple Valued		20,51	0	20,51	20,49	0	20,51	0	20,49	27,99	0
One installment amount	Monthly payment amount of 12 month installment	Financial Properties	Simple Valued		5,7	7,4	6,0	3,9	7,3	4,9	9,3	5,9	10,0	13,7
Colour	Color alternatives of the printer	Physical Appearance	Binary number	5 digit binary number: 1st digit belongs to white, 2nd digit belongs to black, 3rd digit belongs to grey, 4th digit belongs to black-grey, 5th digit belongs to black-white	01000	00010	11000	00001	00001	11001	11000	00001	01001	01000
Volume (mm3)	Volume (mm3) of the printer	Physical Appearance	Simple Valued		31604932	14054656	31447858	23366250	22650584	20624625	16280152	23366250	37716689	46925715
Print Quality (Colour)	The resolution as dots per inch the printer is able to print on non photo paper type in black color	Print Properties	Binary number	5 digit binary number: 1st digit belongs to 600x600, 2nd digit belongs to 1200x600, 3rd digit belongs to 2400x600, 4th digit belongs to 1200x1200, 5th digit belongs to 4800x1200	00001	00000	00001	00001	00000	00001	00000	00001	00001	00000
Print Quality (black)	The resolution as dots per inch the printer is able to print on non photo paper type in color	Print Properties	Simple Valued	5 digit binary number: 1st digit belongs to 600x600, 2nd digit belongs to 1200x600, 3rd digit belongs to 2400x600, 4th digit belongs to 1200x1200, 5th digit belongs to 4800x1200	00001	01000	00001	10000	00100	10000	10000	10000	10000	00010
Print technology	The printing technology type of the printer, eg. thermal inkjet, laser, photo, etc	Print Properties	Type code number	Thermal ink jet: 1 Laser: 2	1	2	1	1	2	1	2	1	1	2
Two side scanning ability	Whether the printer is able to scan both sides of the paper or not	Scan Properties	Simple Boolean		0	0	0	0	0	0	0	0	0	0
Maximum Scan Resolution	Scanning resolution capability as dots per inch	Scan Properties	Binary number	4 digit binary number: 1st digit belongs to 600x600, 2nd digit belongs to 1200x1200, 3rd digit belongs to 600x1200, 4th digit belongs to 1200x2400	0100	0000	0100	0010	0000	0100	0000	0001	0100	1000
Max. doc size for scanning(mm2)	Maximum paper size that the printer can scan	Scan Properties	Simple Valued		64152	0	64152	64161	0	63855	0	64161	76896	76896
Maximum Copy resolution (black)	Copying resolution capability as dots per inch	Scan Properties	Binary number	5 digit binary number: 1st digit belongs to 600x600, 2nd digit belongs to 600x1200, 3rd digit belongs to 1200x1200, 4th digit belongs to 1200x2400, 5th digit belongs to 4800x1200	00100	00000	00100	10000	00000	10000	00000	10000	01000	00000
Maximum Copy resolution (color)	Copying resolution capability as dots per inch	Scan Properties	Binary number	5 digit binary number: 1st digit belongs to 600x600, 2nd digit belongs to 600x1200, 3rd digit belongs to 1200x1200, 4th digit belongs to 1200x2400, 5th digit belongs to 4800x1200	00100	00000	00100	01000	00000	00100	00000	00010	00001	00100
Spare part support after warranty	Whether the printer has service support after warranty period or not	Service Conditions	Simple Boolean		1	1	1	1	1	1	1	1	1	1
On place service	Whether the printer vendor provides service on the buyer's premises or not	Service Conditions	Simple Boolean		1	1	1	1	1	1	1	1	1	1
Call Center	Whether such a center exists or not	Service Conditions	Simple Boolean		1	1	1	1	1	1	1	1	1	1
Warranty Duration	Warranty duration of the printer	Service Conditions	Simple Valued		2	2	2	2	2	2	2	2	2	2
Print Speed (Black A4 ppm)	The number of black and white A4 pages that the printer can print per minute	Speed Properties	Simple Valued		29	16	29	7	20	18	18	8,4	28	22
Print speed (Colour A4 ppm)	The number of colour pages that the printer can print per minute	Speed Properties	Simple Valued		23	0	23	4,8	0	15	0	4,8	22	0
Copy Speed (Black A4)	The number of black and white pages that the printer can print per minute	Speed Properties	Simple Valued		9	0	9	6,3	0	18	0	7,6	28	22
Copy Speed (Colour A4)	The number of colour pages that the printer can copy per minute	Speed Properties	Simple Valued		6,5	0	6,5	2,7	0	15	0	2,6	22	0
Media Feeder Capacity	The number of A4 paper sheets that the printer's mediafeeder can hold	Usage Conveniences	Simple Valued		80	150	80	100	250	80	150	100	100	250
Media Types	The versatility of the media types	Usage Conveniences	Binary number	22 digit binary number: 1st digit belongs to Paper Normal, 2nd digit belongs to Paper Inkjet, 3rd digit belongs to Paper Photo, 4th digit belongs to Envelopes, 5th digit belongs to Transparency, 6th digit belongs to Labels, 7th digit belongs to Cards, 8th digit belongs to Iron Printing, 9th digit belongs to Borderless Media, 10th digit belongs to Thin, 11th digit belongs to Cotton, 12th digit belongs to Recycled, 13th digit belongs to Archive, 14th digit belongs to Coloured, 15th digit belongs to Preprinted, 16th digit belongs to Bond, 17th digit belongs to Thick, 18th digit belongs to Paper High Resolution, 19th digit belongs to Photo Sticker, 20th digit belongs to Paper Laser, 21th digit belongs to Paper Rough, 22th digit belongs to Paper Velum	11111111100 00000000000	10011110011 11111100000	11111111100 00000000000	10110001000 00000011000	10011100000 10001000000	11111111100 00000000000	10111110000 00000000111	10110001000 00000011000	11111011000 00000000000	10011110000 10110000000
Media Sizes	The types of paper sizes that the printer can print on	Usage Conveniences	Binary number	29 digit binary number: 1st digit belongs to A4, 2nd digit belongs to A5, 3rd digit belongs to A6, 4th digit belongs to B5, 5th digit belongs to B6, 6th digit belongs to C5, 7th digit belongs to C6, 8th digit belongs to nvelope DL, 9th digit belongs to Envelope No 9, 10th digit belongs to Envelope No 10, 11th digit belongs to Envelope C5, 12th digit belongs to Envelope B5, 13th digit belongs to Envelope Monarch, 14th digit belongs to Legal, 15th digit belongs to Executive, 16th digit belongs to Efolio, 17th digit belongs to Folio, 18th digit belongs to Letter, 19th digit belongs to 300x100, 20th digit belongs to 130x180, 21th digit belongs to 100x150 (4x6) 22th digit belongs to 77x127 (3x5), 23th digit belongs to 215x762, 24th digit belongs to 215x279 (8.5x11), 25th digit belongs to 70x216, 26th digit belongs to 116x356, 27th digit belongs to 203x254 (8x10), 28th digit belongs to 127x178 (5x7), 29th digit belongs to 147x211	11110011000 00000001111 1000000	11010101110 01011100001 0000000	11010000000 00110010011 0100000	11010001010 00100010010 0000110	11111001000 00110010000 0011000	11110011000 00000011111 0000000	11110101000 10000010000 0000101	11010001010 00100010010 0000110	11010011000 00000000110 0000000	11110111010 00111110001 0000000
Two side printing option	Whether the printer has two sided printing option or not	Usage Conveniences	Simple Boolean		1	0	1	1	1	1	1	1	0	1
Screen	Whether the printer has a screen or not	Usage Conveniences	Simple Boolean		0	0	1	1	0	0	0	1	1	1
Fax ability		Usage Conveniences	Simple Boolean		0	0	0	0	0	0	0	0	1	0
Weight	Weight of the printer	Usage Conveniences	Simple Valued		4,618	4,95	4,6	5,6	6,5	4,9	5,2	5,5	5,9	10,59

Table-4: Differences of property values between HP C4680 and alternative printers

Property name	Samsung ML1660	HP Photosmart	Canon PIXMA	Brother HL-2040	HP Deskjet F2420	HP laserjet ProP1102	Canon Pixima	HP Officejet J4500	Samsung SCX4600
Interface and connectivity	0	1	0	1	0	0	0	1	0
Memory card options	5	0	5	5	5	5	1	5	5
Power	230,00	0,00	9,00	430,00	2,80	340,00	8,00	10,40	666,00
Memory	56,00	0,00	62,00	56,00	62,00	62,00	62,00	0,00	0,00
Third party cartridge refill	0	1	1	1	1	1	1	1	0
Price (USD exld. Wat)	20,00	3,50	22,00	18,50	9,50	43,00	2,00	51,40	96,40
Initial cartridge level percentage	0,00	0,00	0,50	0,50	0,50	0,00	0,50	0,00	0,00
Cartridge cost black (USD excld.	52,50	0,00	2,50	40,50	0,00	61,01	2,50	0,00	71,50
Cartridge cost color (USD excld.	20,51	0,00	0,02	20,51	0,00	20,51	0,02	7,48	20,51
One installment amount.	1,67	0,29	1,83	1,54	0,79	3,58	0,17	4,28	8,03
Colour.	2	1	2	2	2	1	2	1	0
Volume (mm3)	17550275,73	157073,73	8238681,73	8954347,73	10980306,73	15324779,73	8238681,73	6111757,44	15320783,47
Print Quality (Colour)	1	0	0	1	0	1	0	0	1
Print Quality (black)	2	0	2	2	2	2	2	2	2
Print technology	0	1	1	0	1	0	1	1	0
Two side scanning ability	0	0	0	0	0	0	0	0	0
Maximum Scan Resolution	1	0	2	1	0	1	2	0	2
Max. doc size for scanning(mm2)	64152,00	0,00	9,00	64152,00	297,00	64152,00	9,00	12744,00	12744,00
Maximum Copy resolution (black)	1	0	2	1	2	1	2	2	1
Maximum Copy resolution (color)	1	0	2	1	0	1	2	2	0
Spare part support after	1	1	1	1	1	1	1	1	1
On place service	1	1	1	1	1	1	1	1	1
Call Center	1	1	1	1	1	1	1	1	1
Warranty Duration.	0	0	0	0	0	0	0	0	0
Print Speed (Black A4 ppm)	13,00	0,00	22,00	9,00	11,00	11,00	20,60	1,00	7,00
Print speed (Colour A4 ppm)	23,00	0,00	18,20	23,00	8,00	23,00	18,20	1,00	23,00
Copy Speed (Black A4)	9,00	0,00	2,70	9,00	9,00	9,00	1,40	19,00	13,00
Copy Speed (Colour A4)	6,50	0,00	3,80	6,50	8,50	6,50	3,90	15,50	6,50
Media Feeder Capacity	70,00	0,00	20,00	170,00	0,00	70,00	20,00	20,00	170,00
Media Types	12	0	8	7	0	6	6	2	7
Media Sizes	12	10	11	12	2	11	11	4	11
Two side printing option	0	1	1	1	1	1	1	0	1
Screen	0	1	1	0	0	0	1	1	1
Fax ability.	0	0	0	0	0	0	0	1	0
Weight.	0,33	0,02	0,98	1,88	0,28	0,58	0,88	1,28	5,97

Table-5: Partial Fuzzy membership function values of each printer with respect to properties: $\mu_x(\mathbf{r}_{yijk})$

Property name	Samsung ML1660	HP Photosmart C4780	Canon PIXMA MP250	Brother HL-2040	HP Deskjet F2420	HP laserjet ProP1102	Canon Pixima MP490	HP Officejet J4500	Samsung SCX4600
Brand name.	0,00	1,00	0,00	0,00	1,00	1,00	0,00	1,00	0,00
Interface and connectivity	0,00	1,00	0,00	1,00	0,00	0,00	0,00	1,00	0,00
Memory card options	1,00	0,80	1,00	1,00	1,00	1,00	0,84	1,00	1,00
Power	0,65	1,00	0,99	0,35	1,00	0,49	0,99	0,98	0,00
Memory	0,10	1,00	0,00	0,10	0,00	0,00	0,00	1,00	1,00
Third party cartridge refill options	0,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00
Price (USD exld. Wat)	0,79	0,96	0,77	0,81	0,90	0,55	0,98	0,47	0,00
Initial cartridge level percentage	1,00	1,00	0,00	0,00	0,00	1,00	0,00	1,00	1,00
Cartridge cost black (USD excld. Wat)	0,27	1,00	0,97	0,43	1,00	0,15	0,97	1,00	0,00
Cartridge cost color (USD excld. Wat)	0,00	1,00	1,00	0,00	1,00	0,00	1,00	0,64	0,00
One installment amount.	0,79	0,96	0,77	0,81	0,90	0,55	0,98	0,46	0,00
Colour.	0,00	0,50	0,00	0,00	0,00	0,50	0,00	0,50	1,00
Volume (mm3)	0,00	0,99	0,53	0,49	0,37	0,13	0,53	0,65	0,13
Print Quality (Colour)	0,00	1,00	1,00	0,00	1,00	0,00	1,00	1,00	0,00
Print Quality (black)	0,00	1,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Print technology	0,00	1,00	1,00	0,00	1,00	0,00	1,00	1,00	0,00
Two side scanning ability	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Maximum Scan Resolution	0,50	1,00	0,00	0,50	1,00	0,50	0,00	1,00	0,00
Max. doc size for scanning(mm2)	0,00	1,00	1,00	0,00	1,00	0,00	1,00	0,80	0,80
Maximum Copy resolution (black)	0,50	1,00	0,00	0,50	0,00	0,50	0,00	0,00	0,50
Maximum Copy resolution (color)	0,50	1,00	0,00	0,50	1,00	0,50	0,00	0,00	1,00
Spare part support after warranty	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
On place service	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Call Center	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Warranty Duration.	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Print Speed (Black A4 ppm)	0,41	1,00	0,00	0,59	0,50	0,50	0,06	0,95	0,68
Print speed (Colour A4 ppm)	0,00	1,00	0,21	0,00	0,65	0,00	0,21	0,96	0,00
Copy Speed (Black A4)	0,53	1,00	0,86	0,53	0,53	0,53	0,93	0,00	0,32
Copy Speed (Colour A4)	0,58	1,00	0,75	0,58	0,45	0,58	0,75	0,00	0,58
Media Feeder Capacity	0,59	1,00	0,88	0,00	1,00	0,59	0,88	0,88	0,00
Media Types	0,00	1,00	0,33	0,42	1,00	0,50	0,50	0,83	0,42
Media Sizes	0,00	0,17	0,08	0,00	0,83	0,08	0,08	0,67	0,08
Two side printing option	0,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
Screen	1,00	0,00	0,00	1,00	1,00	1,00	0,00	0,00	0,00
Fax ability.	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,00	1,00
Weight.	0,94	1,00	0,84	0,68	0,95	0,90	0,85	0,79	0,00

Table-6: Final similarity scores of each alternative printer with respect to HP C4680

Property name	Samsung ML1660	HP Photosmart C4780	Canon PIXMA MP250	Brother HL-2040	HP Deskjet F2420	HP laserjet ProP1102	Canon Pixima MP490	HP Officejet J4500	Samsung SCX4600
Brand name.	0,000	0,024	0,000	0,000	0,024	0,024	0,000	0,024	0,000
Interface and connectivity	0,000	0,051	0,000	0,051	0,000	0,000	0,000	0,051	0,000
Memory card options	0,030	0,024	0,030	0,030	0,030	0,030	0,025	0,030	0,030
Power	0,003	0,004	0,004	0,001	0,004	0,002	0,004	0,004	0,000
Memory	0,001	0,007	0,000	0,001	0,000	0,000	0,000	0,007	0,007
Third party cartridge refill options	0,000	0,064	0,064	0,064	0,064	0,064	0,064	0,064	0,000
Price (USD exld. Wat)	0,028	0,034	0,027	0,029	0,032	0,020	0,035	0,017	0,000
Initial cartridge level percentage	0,005	0,005	0,000	0,000	0,000	0,005	0,000	0,005	0,005
Cartridge cost black (USD excld. Wat)	0,004	0,013	0,013	0,006	0,013	0,002	0,013	0,013	0,000
Cartridge cost color (USD excld. Wat)	0,000	0,012	0,012	0,000	0,012	0,000	0,012	0,008	0,000
One installment amount.	0,028	0,034	0,027	0,029	0,032	0,020	0,035	0,016	0,000
Colour.	0,000	0,001	0,000	0,000	0,000	0,001	0,000	0,001	0,003
Volume (mm3)	0,000	0,013	0,007	0,007	0,005	0,002	0,007	0,009	0,002
Print Quality (Colour)	0,000	0,037	0,037	0,000	0,037	0,000	0,037	0,037	0,000
Print Quality (black)	0,000	0,037	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Print technology	0,000	0,113	0,113	0,000	0,113	0,000	0,113	0,113	0,000
Two side scanning ability	0,011	0,011	0,011	0,011	0,011	0,011	0,011	0,011	0,011
Maximum Scan Resolution	0,023	0,046	0,000	0,023	0,046	0,023	0,000	0,046	0,000
Max. doc size for scanning(mm2)	0,000	0,091	0,091	0,000	0,090	0,000	0,091	0,073	0,073
Maximum Copy resolution (black)	0,010	0,019	0,000	0,010	0,000	0,010	0,000	0,000	0,010
Maximum Copy resolution (color)	0,010	0,019	0,000	0,010	0,019	0,010	0,000	0,000	0,019
Spare part support after warranty	0,004	0,004	0,004	0,004	0,004	0,004	0,004	0,004	0,004
On place service	0,007	0,007	0,007	0,007	0,007	0,007	0,007	0,007	0,007
Call Center	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002
Warranty Duration.	0,012	0,012	0,012	0,012	0,012	0,012	0,012	0,012	0,012
Print Speed (Black A4 ppm)	0,024	0,058	0,000	0,034	0,029	0,029	0,004	0,056	0,040
Print speed (Colour A4 ppm)	0,000	0,050	0,011	0,000	0,033	0,000	0,011	0,048	0,000
Copy Speed (Black A4)	0,008	0,016	0,013	0,008	0,008	0,008	0,014	0,000	0,005
Copy Speed (Colour A4)	0,009	0,016	0,012	0,009	0,007	0,009	0,012	0,000	0,009
Media Feeder Capacity	0,004	0,007	0,006	0,000	0,007	0,004	0,006	0,006	0,000
Media Types	0,000	0,020	0,007	0,008	0,020	0,010	0,010	0,016	0,008
Media Sizes	0,000	0,003	0,001	0,000	0,014	0,001	0,001	0,011	0,001
Two side printing option	0,000	0,020	0,020	0,020	0,020	0,020	0,020	0,000	0,020
Screen	0,015	0,000	0,000	0,015	0,015	0,015	0,000	0,000	0,000
Fax ability.	0,081	0,081	0,081	0,081	0,081	0,081	0,081	0,000	0,081
Weight.	0,006	0,007	0,006	0,005	0,006	0,006	0,006	0,005	0,000
Final Similarity Score	**0,323**	**0,961**	**0,616**	**0,474**	**0,797**	**0,431**	**0,634**	**0,695**	**0,347**