International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 1 Issue 9, November- 2012

Fuzzy Clustering For The Problems Faced By IT Professionals Due To Work Pressure In Their IT Field

M.Geethalakshmi*, A.Rajkumar, N.Joseparvinpraveena*****
KCG College of Technology
****Hindustan University**

ABSTRACT

In our day-to-day life the growing organization such as IT Companies plays an important role in their field. Especially IT Professionals, they are facing so many problems both generally as well health wise. Because of their work pressure they are facing lot of problems. For these circumstances, twenty attributes have been chosen in this study and the problems are taken in a 10-point rating scale regarding the IT Problems in the level of frustration/stress at the chosen attributes of the problems faced by the IT Professionals in Chennai city. The average rating scale is using to cluster the attributes using Fuzzy c-means clustering and classify them as low, medium and high level of frustration/stress according to their problems and it can be classified proactively to ensure that the non-occurrence of problems, the advantage of fuzzy clustering is that attributes may be part of more than one cluster to a varying degree and this gives a better picture about the attributes of IT Problems.

Keywords – C-means Clustering, Degree, Mid-Value, Stress, IT Professionals.

1. INTRODUCTION

Computer, a hallmark of technological advancement has ushered in a new genre of occupational health problem, i.e. of computer related health problems. India being the forerunner in the cyber world the occupational health personals is slowly awakening to this group of modern occupational diseases, which are slowly taking its roots among the information technology (IT) professionals. There is an urgent need to understand the dynamics of these problems and prevent it from assuming epidemic proportions. IT Professionals facing so many problems in their day-to-day life. Both general and health related problems which lead to the level of frustration/stress. Among the survey from 100 IT Professionals, the most commonly occurring problems are listed and that are taken as the major attributes. This study aims to capture the IT Professionals problem about the level of frustration/stress using the attributes and classify that attributes based on fuzzy c-means clustering. Fuzzy clustering is more appropriate for perceptions based data since perceptions are always a matter of varying degree. The organization of this article is as follows: Choosing the attributes in methodology section, Basic notion of clustering and fuzzy clustering in next section, Results and Discussion in the fourth section and finally we arrive at the conclusion and scope for our future study.

2. METHODOLOGY

In this methodology, the classification of IT Problems due to work pressure is done based on the IT Professionals perception to select the attributes of IT Problems. The following twenty attributes of Problems is chosen by interviewing the IT Professionals in Chennai at the different ages.

1. Divorce
2. Misunderstandings
3. Arguments with partners
4. Working in shifts
5. Insomnia(Inability to sleep)
6. Mental Stress
7. Vision Problem
8. Less Immunity feeling sick always
9. Staying away from Native
10. Frustration
11. Depression
12. Obesity
13. Ulcer/Acidity
14. Joint Pain
15. Back Pain
16. Shoulder Pain by sitting in same place
17. Burning Eyes
18. Loss of Energy/ Loss of Money on Travelling
19. Head Ache
20. Tension

Three attributes that defines the best characteristic of each segment have been selected to be rated by respondents. Fuzzy c-means clustering is done using the algorithm (4.1).

2. PRELIMINARIES

(a)HARD CLUSTERING

In **Hard Clustering** we make a hard partition of the data set **Z**. In other words, we divide them into $c \geq 2$ clusters. With a partition, we mean that

$$\bigcup_{i=1}^{c} A_i = Z$$

and $A_i \cap A_j = \phi$, $\forall\, i \neq j$ ..... (1)

Also, none of the sets, A_i may be empty. To indicate a portioning, we make use of **membership functions** $\mu_k(x)$. If $\mu_k(x) = 1$, then object x is in cluster k. Based on the membership functions, we can assemble the **Partition Matrix U**, of which $\mu_k(x)$ are the elements. Finally there is a rule that $\forall$x,

$$\sum_{i=1}^{c} \mu_k \ (x) = 1 \quad \forall x \quad \ldots\ldots\ldots\ldots\ (2)$$

In other words, every object is only part of one cluster.

(b)FUZZY CLUSTERING

Hard clustering has a downside. When an object roughly falls between two clusters A_i and A_j, it has to be put into one of these clusters. Also, outliers have to be put in some cluster. This is undesirable. But it can be fixed by fuzzy clustering.

In **Fuzzy clustering**, we make a **Fuzzy partition** of the data. Now, the membership function $\mu_k(x)$ can be any value between 0 and 1. This means that an object z_k can be for 0.2 parts in A_i and for 0.8 parts in A_j. However, requirement (2) still applies. So, the sum of the membership functions still has to be 1. The set of all fuzzy partitions that can be formed in this way is denoted by M_{fc}. Fuzzy partitioning again has a downside. When we have an **outlier** in the data (being an object that doesn't really belong to any cluster), we still have to assign it to clusters. That is, the sum of its membership functions still must equal one.

(c) FUZZY C-MEANS CLUSTERING

In fuzzy clustering, each point has a degree of belonging to clusters, as in fuzzy logic, rather than belonging completely to just one cluster. Thus, points on the edge of a cluster, may be in a cluster to a lesser degree than points in the center of cluster for each point x there is no coefficient giving the degree of belonging in the k^{th} cluster $\mu_k(x) = 1$. Usually, the sum of those coefficients is defined to be 1.

$$\sum_{n=1}^{mm\ cluster} \mu_k \quad = 1 \quad \forall x \ldots\ldots\ldots \ (3)$$

With fuzzy c-means, the centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster

$$\text{Center}_k = \frac{\sum_x \mu_k(x)^m x}{\sum_x \mu_k(x)^m} \quad \ldots\ldots\ldots\ldots\ldots\ (4)$$

The degree of belonging is related to the inverse of the distance to the cluster

$$\mu_k(x) = \frac{1}{d(center_k, x)} \ldots\ldots\ldots\ldots\ldots\ (5)$$

then the coefficients are normalized and fuzzy field with a real parameter $m > 1$ so that their sum is 1. So

$$\mu_k(x) = \frac{1}{\sum_j \left(\frac{d(center_k, x)}{d(center_j, x)}\right)^{2/(m-1)}} \ldots\ldots\ldots\ldots (6)$$

For m equal to 2, this is equivalent to normalizing the coefficient linearly to make their sum 1. When m is close to 1, then cluster center closes to the point is given much more weight than the others, and the algorithm is similar to k-means.

4. RESULTS AND DISCUSSION

We have interviewed 100 IT Professionals in Chennai city to find what are all the problems they are facing in IT field, for that twenty attributes of has been chosen and the respondents had related the attributes of problems engendered high level frustration on a 10-point rating scale and the results of the average rating scale and the results of the average rating are shown in **Fig.1** 5^{th} and 6^{th} attributes is rated highest by the respondents with an average rating of 8.1 and 8.9 respectively on a 10-point scale. This means that, when we compare to other attributes 6^{th} and 5^{th} engendered high level frustration/stress and the 18^{th} attribute rating average 2.5 which engendered low level of frustration/stress.

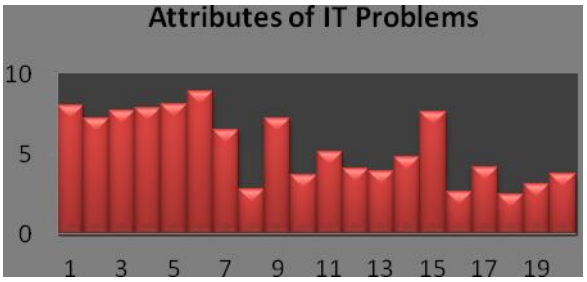

Fig.1 Mean rating of frustration/stress level

The ratings and the Standard Deviation of the attributes of the problems faced by IT Professionals engendered high level frustration have been subjected to fuzzy c-means clustering using the algorithm:1 and the following results shown in Table :1 have been obtained for a 3-cluster combination. The first cluster comprises of the attributes with average rating from 2.5 to 5.5 with a mid value 4. The second cluster range is from 3.5 to 8.5 with a mid valued 6 and the third cluster has a range of 7.5 to 10 with a mid value 8.75.

The first cluster range indicates the problems of IT Professionals which engendered LOW level frustration. Second and third clusters range shows that MODERATE and HIGH level of frustration/Stress respectively. There is Over-lapping ranges as in characteristic of a fuzzy based cluster.

Table: 1
3-Cluster Range of Level of Frustration/Stress

	Cluster 1	Cluster 2	Cluster 3
Range	2.5-5.5	3.5-8.5	7.5-10
Mid- Value	4	6	8.75
Classification	LOW	MODERATE	HIGH

4.1 ALGORITHM TO FIND A MEMBERSHIP VALUES FOR THE ATTRIBUTES

Step 1: Start
Step 2: Fix, the values of 20 attributes on a 10-point rating scale in a set D (say)
Step 3: Fix the cluster, which is defined as Cluster 1 = LOW, whose range beginning with 2.5 (**bv1**) End with 5.5(**ev1**). Cluster 2 = MODERATE, whose range beginning with 3.5 (**bv2**) end with 8.5(**ev2**). Cluster 3 = HIGH, whose range beginning with 7.5 (**bv3**) end with 10 (**ev3**).
Step 4: Choose en element x in D
Step 5: If $x <$ **ev1**, Go to *Step 6*, else Go to *Step8*
Step 6: If $x >$ **bv2**, then x **lies in cluster 1 and cluster 2** whose membership value is defined as $\mu_k(x) =$ **ev1-**x**:** x**- bv2**, Go to *Step 12*, else Go to *Step7*.
Step7: Then x **lies in cluster 1** only, the membership value is $\mu_k(x) =$**1** Go to *Step 12*
Step 8: If $x <$ **ev2** Go to *Step 9*, else Go to *Step 11*
Step 9: If $x >$ **bv3**, then x **lies in cluster 2 and cluster 3**, whose membership value is defined as $\mu_k(x) =$ **ev2-**x**:** x**-bv3**, Go to step 12, else Go to *Step10*
Step 10: Then x **lies in cluster 2** only, the membership value is $\mu_k(x) =$**1** else Go to *Step11*
Step 11: Then x **lies in cluster 3** only, the membership value is $\mu_k(x) =$**1**
Step 12: Go to *Step 4*, until all the values in D have been checked
Step 13: Stop

Here **'bv'** denotes the beginning value and '**ev**' denotes the ending the value.

Degree of membership of the attributes of IT Problems is found using the above algorithm is shown in Table: II. Attributes 8,16,18,19 with a mean rating 2.8, 2.6, 2.5, 3.1 is entirely (100 %) with a membership value of 1in cluster 1. (i.e.) frustration/ stress level is LOW.

Attributes 2, 7, 9 with a mean rating 7.2, 6.9, 7.2 is entirely (100 %) with a membership value of 1in cluster 2. (i.e.) frustration/ stress level is MODERATE.

Attribute 6 with a mean rating 8.9, is entirely (100 %) with a membership value of 1in cluster 3. (i.e.) frustration/ stress level is HIGH.

Attributes 1, 3, 4, 5 &15 with a mean ratings 8,7.7,7.9,8.1,&7.6 belongs to 50 % in cluster 2 and50 % in cluster 3, 80 % in cluster 2 and 20 % in cluster 3, 60 % in cluster 2 and 40 % in cluster 3, 40 % in cluster 2 and 60 % in cluster 3, 90 % in cluster 2 and 10 % in cluster 3. (i.e.)frustration/ stress level is between MODERATE and HIGH.

Attributes 10, 11, 12, 13, 14, 17 &20 with a mean ratings 3.7,5.1,4.1,3.9,4.8,4.2 &3.8 belongs to 90 % in cluster 1 and10 % in cluster 2, 20 % in cluster 1 and 80 % in cluster 2, 70 % in cluster 1 and 30 % in cluster 2, 80 % in cluster 1 and 20 % in cluster 2, 35 % in cluster 1 and 65 % in cluster 2, 65 % in cluster 1 and 35 % in cluster 2, 85% in cluster 1 and 15% in

cluster 2. (i.e.) frustration/ stress level is between LOW and MODERATE.

Table: II
Degree of Membership of the Attributes

S.No	Mean	LOW	MODERATE	HIGH
1	8	0	0.50	0.50
2	7.20	0	1	0
3	7.7	0	0.80	0.20
4	7.9	0	0.60	0.40
5	8.1	0	0.40	0.60
6	8.9	0	0	1
7	6.5	0	1	0
8	2.8	1	0	0
9	7.2	0	1	0
10	3.7	0.90	0.10	0
11	5.1	0.20	0.80	0
12	4.1	0.70	0.30	0
13	3.9	0.80	0.20	0
14	4.8	0.35	0.65	0
15	7.6	0	0.90	0.10
16	2.6	1	0	0
17	4.2	0.65	0.35	0
18	2.5	1	0	0
19	3.1	1	0	0
20	3.8	0.85	0.15	0

5. CONCLUSION

In Crisp Clustering, where a attributes is a member of one cluster only, the fuzzy clustering process permits a attributes to be a member of more than one cluster although to a varying degree. This helps us to find out where IT Professional gets more frustrated and which stimulate their frustration/stress so that we can give remedies. From our study we conclude that the IT professionals insisted on knowing where they were always; if they have arguments with partners, working in shifts, Insomnia, having mental stress which stimulates the frustration/stress thought in them very high. In this article, we observed that the effect of mental stress, Insomnia and divorce having high mean rating and this will not only leads to frustration/stress thought but also divorce between the couple which will be our main scope of study in the future to analyze the causes of divorce.

REFERENCES

[1] A.J. Zolotor, A.C. Denham, A. Weil, "*Intimate partner Violence*", Prim Care 2009.

[2] A. Kaufmann, "*Introduction to the Theory of Fuzzy Subsets*", Academic Press, INC. (LONDON) LTD, 1975.

[3] A. Resolution/RES/48/104, *Declaration on the elimination of violence against women*, adopted by the UN General Assembly, 20 December 1993.

[4] B. Kosko, "*Hidden pattern in combined and Adaptive Knowledge Networks*" Proc. Of the First, *IEE International Conference on Neural Networks* (ICNN-86(1988) 377-398).

[5] B. Kosko, "*Neural Networks and Fuzzy systems: A Dynamical System Approach to Machine Intelligence*", Prentice Hall of India, 1997.

[6] C. Gacria-Moreno, HA. Jansen, M. Elsberg, L. Hesise, CH. Waltts, "*Definitions Questionnaire development In. Who multi-country study on women's health and domestic violence against women*", Geneva World Health Organization, 2005.

[7] D. Varma, PS. Chandra, T. Thomas, MP. Carey, *Intimate partner violence and sexual coercion among pregnant women in India: Relationship with depression and post traumatic stress disorder.* J Affect Disord 2007; 102:227-35.

[8] H.J. Zimmermann, "*Fuzzy Set Theory and its application*", Fourth Edition Springer 2011.

[9] J. Klir George/Bo Yuan, "*Fuzzy sets and Fuzzy Logic : Theory and Applications*", Prentice Hall of India.

[10] M.A. Strauss, "*Conflict Tactics Scales. In: NA. Jackson, editor Encyclopedia of Domestic violence*", 2007.

[11] S. Kumar, L. Jeyaseelan, S. Suresh, RC, Ahuja, "*Domestic violence and its mental health correlates in Indian women*", Br J Psychiatry 2005.

[12] Sa. Falsetti, "*Screening and responding to family and intimate partner violence in the primary care setting*", Prim Care 2007.

[13] J. Gunter, "Intimate Partner Violence. Obset Gynecol Clin North Am 2007; 34:367-88.

[14] WHO multi-country study on women's health and domestic violence against women: summary report of Indian results on prevalence, health outcomes and Women's responses, Geneva: World Health Organization; 2005, p-15-7.

[15] Fuzzy Clustering of the Domestic Violence for the Degree of Suicide thought based on Married Women Perception by A.Victor Devadoss, A.Felix, 2012, IJERA p: 1770-1773