International Journal of Intelligent Information Technologies
Volume 12 • Issue 2 • April-June 2016

Covering Rough Clustering Approach for Unstructured Activity Analysis

Prabhavathy Panneer, School of Information Technology and Engineering, VIT University, Vellore, India

B.K. Tripathy, School of Computing Science and Engineering, VIT University, Vellore, India

ABSTRACT

Several tasks under human activities need to be performed in a sequence of navigation and manipulation of objects. In several applications of human activities like robotics monitoring plays an important role. So, in these applications, processing of sequential data is of utmost importance. Because of the presence of imprecision intelligent clustering approaches using fuzzy or rough set techniques play a major role. The basic rough sets which are defined by using equivalence relations is less useful because of their scarcity in real life scenarios. As a result, covering based rough sets have been introduced which are more general and applicable to real world problems. In this paper, covering rough set based clustering approach is introduced and studied using refined first type of covering based rough sets. Through experimental analysis,illustrated the efficiency of proposed algorithm and provided a comparative analysis of this algorithm with other existing algorithms.

KEYWORDS

Acitvity, Clustering, Covering, Relative Similarity, Robotics, Rough Set, Sequential, Similarity

1. INTRODUCTION

The location and configuration of the objects involved to often vary in erratic ways in unstructured human environments. This requires processing sequence of activities that is robust plus ñexible in an uncertain environment(Hema Koppula & Ashutosh Saxena, 2015). Sequential pattern is a non-empty ordered list of items. The clustering of sequential patterns is a relevant and scalable solution for behavior modeling (Lingras,P. & West, C.,2004). The original rough set theory proposed by Pawlak (Pawlak, Z., 1982) is based upon equivalence relations defined over a universe. One of the recent generalizations of this notion is covering based rough sets, introduced by Zakowski (Zakowski, W,1983). A cover is a generalization of the notion of partition where components can have non empty intersection. The covering based rough sets are models with promising potential for the field of data mining. Various types of covering based rough sets and their properties have been discussed by many researchers(Pawlak, Z., 1991; Tripathy, B.K et al ., 2008; Tripathy, B.K & Tripathy HK.,2009) for various applications.

Clustering techniques have been effectively applied to a wide range of engineering and scientific disciplines such as pattern recognition, machine learning, psychology, biology, medicine, computer vision, communications, and remote sensing. Clustering is categorized as hard or soft in nature. Soft clusters may have fuzzy or rough boundaries. A number of clustering algorithms have been proposed to suit different requirements. Rough clustering can help researchers to discover overlapping clusters in many applications such as sequential mining and text mining. Clustering Sequential data is one of the vital research tasks. Several tasks in human environments need performing a sequence of navigation

DOI: 10.4018/IJIIT.2016040101

and manipulation steps involving objects. Human activity monitoring and detections are prominent for a number of applications. Sequential data processing is the significant research tasks. Soft clustering approach such as rough set and fuzzy performs a major role in sequential data applications.

This paper proposed refined first type covering based rough set approach to cluster sequential data. The remainder of this paper is organized as follows. The related work is discussed in section2. In Section 3, a brief introduction to the S^3M similarity measure for sequential data is presented. Section 4 discusses covering clustering approach using first type coverings based rough set. Section 5 investigates the experimental and result analysis of the proposed algorithm and concluded the work done in this paper in section 6.

2. RELATED WORKS

Data are generated and collected using sensors, cameras and smart devices. The generated data create vagueness, incompleteness, and granularity in an information system which gives unreliable solution in the data analysis. In our day to day approach, several traditional tools are used for formal modeling, computing and reasoning of data, which are basically crisp and deterministic in nature. Real situations are very often not crisp and deterministic, and they cannot be described precisely. Rough Sets is one of the excellent modeling tool introduced by Pawlak (Pawlak, Z.,1982,1991) which is used to study the imprecise data in the information system. Applying rough sets in real-time information systems leads to some restriction because of the equivalence relation unless the clustering of problem appears to hold true for equivalence relation.The equivalence relations of rough sets were extended to generalized binary relations in several directions. Similarly, partitions of the universes used to define rough sets were extended to coverings. A type of generalized rough sets based on covering and the relationship between this type of covering- based rough sets (Zhanhong Shi & Zengtai Gong, 2010) and the generalized rough sets based on binary relation were studied. Clusters can be hard or soft in nature. In soft clustering, an object may be a member of two or more clusters. Soft clusters may have fuzzy or rough boundaries. Web mining is one such area where overlapping clusters are required. Generally, clustering algorithms make use of either distance functions or similarity functions for comparing pairs of sequences.

Data is in the form of sequences in variety of domains like robotics, healthcare, education, bioinformatics, telecommunications, log analysis, anomaly detection, etc. Edit distance and LCS are the two major approaches followed so far to measure the similarities (Hassan Saneifar et al ., 2008) in sequential data. A S2MP (Similarity Measure for Sequential Patterns) similarity measure was defined for computing the similarity between of sequential patterns, which take the characteristics and the semantics of sequential patterns into account. The clustered sequential patterns can, for instance, be used to analyze user activity in unstructured environment. Categorization of Clustering algorithms has been done using diverse taxonomies based on different important issues such as algorithmic structure, nature of clusters formed, use of feature sets, etc(Jain,et al.,1999) . The clustering of sequential patterns is a relevant and scalable solution for behavior modeling (Hassan Saneifar et al., 2008). Clusters may be of hard or soft nature. In hard clusters, the elements which are similar to each other are placed in the same cluster. The elements whose natures differ with each other drastically are placed in different clusters.

The advantage of using rough sets is that, unlike other techniques, rough set theory does not require any prior information about the data such as apriori probability in statistics and a membership function in fuzzy set theory. The rough sets and fuzzy sets concepts were integrated (Bozkir and Ebru, 2013; Krishnasamy et al.,2014) and applied in information system. Attribute weighted fuzzy clustering has become a very active area of research and interval number has been introduced for attribute weighting in the weighted fuzzy c-means (WFCM) clustering approach (Chatzis, 2011;Zhang et al.,2014). The extensive survey of the significant extensions and derivatives of soft clustering approaches have been studied (Peters et al., 2013; Li et al.,2014).

3. SEQUENTIAL DATA AND SIMILARITY MEASURE

Data which occur one after another with respect to their positions but not necessarily related to each other is termed as sequential data. A sequence is an ordered set of items. Typically, a sequence is denoted as $S = \{a_1, a_2, ..., a_n\}$ where a_1 ; a_2 ; . . . ; a_n are the ordered items in the sequence S. The number of items present in a sequence is defined as the length of the sequence, and it is denoted as $|S|$ called the cardinality of S. In order to find patterns in sequences, it is not only necessary to look at the items contained in the sequences but also it is necessary to note the order of their occurrence.

A new measure, called sequence and set similarity measure (S^3M) was introduced in the clustering domain (Pradeep Kumar, P et al., 2008). The S^3M measure is subdivided into two parts; one that quantifies the composition of the sequence (set similarity) and the other that quantifies the sequential nature (sequence similarity).Sequence similarity quantifies the amount of similarity with respect to the occurrence of item sets within two sequences. Length of the longest common subsequence (LLCS) (Bergroth, L. et al., 2000) with respect to the length of the longest sequence gives the sequence similarity aspect between two sequences. For instance, assume two sequences $S_1$and S_2, then their sequence similarity measure is shown in (1).

$$\text{SeqSim}(S_1,S_2)=\frac{\text{LLCS}(S_1,S_2)}{\max(|S_1|,|S_2|)} \tag{1}$$

Set similarity (Jaccard similarity measure) produces the ratio of the number of common item sets and the number of unique item sets in two sequences. Thus, for two sequences $S_1$and S_2, the set similarity measure is shown in (2).

$$\text{SeqSim}(S_1,S_2)=\frac{|S_1 \cap S_2|}{|S_1 \cup S_2|} \tag{2}$$

Combine set similarity and sequence similarity components into one function in order to take care of both the content as well as position based similarity aspects. Thus, S^3M measure for two sequences S_1 and S_2 is represented by (3).

$$S^3M(S_1,S_2) = p * \frac{\text{LLCS}(S_1,S_2)}{\max(|S_1|,|S_2|)} + q * \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|)} \tag{3}$$

Here, $p + q = 1$ and $p, q \geq 0$ .Infact, p and q provide the relative weights to be given for an order of occurrence (sequence similarity) and to the content (set similarity) respectively. The relative weights are supplied by user between 0 and 1. Even dynamic programming approach can be handled using the LLCS between two sequences (Jain et al.,1999).

4. COVERING BASED ROUGH CLUSTERING

Generalization of rough set theory can be done mainly in three ways, namely, set-theoretic framework using non-equivalence binary relations; granule based definition using coverings and subsystems based definition using other subsystems. Partitions used in rough sets are extended to covers in covering based rough sets (Yiyu Yao & Bingxue Yao,2012) .Here, the members of covers are overlapped where as in partitions members are disjoint in nature. Covering based rough sets (Tripathy, B.K. & Mitra,A.,2009) are classified into various types, such as first type, second type and third type covering based roug sets (Zhu, W. & Fei-Yue Wang,2007). In this paper, refined first type Covering based rough set is used for intelligent sequential clustering approach. The basic definitions and types of covering based rough set are as follows.

Definition 1: Let U be a universe of discourse, $C = \{X \mid X \subseteq U\}$ a family of subsets of U, If no element of C is empty, and $\bigcup_{X \in C} X = U$, then C is called a covering of U. The ordered pair (U,C) a covering approximation space[14].

Definition 2: Let (U,C) be a covering approximation space, $x \in U$, then set family $\mathrm{Md}(x) = \{K \in C \mid x \in K \wedge (\forall S \in C \wedge x \in S \wedge S \subseteq K \Rightarrow K = S)\}$ is called the minimal description of x

Definition 3: Let(U,C) be a covering approximation space. For any $X \subseteq U$, the first type covering lower and upper approximations of X with respect to covering approximation space(U,C) are defined as follows:

$$\underline{FC}(X) = \bigcup\{K \in C \mid K \subseteq X\} \quad \overline{FC}(X) = \underline{FC}(X) \cup \{Md(x) \mid x \in X - \underline{FC}(X)\}$$

Where $\underline{FC}(X)$ denotes first type covering lower approximation and $\overline{FC}(X)$ denotes first type covering upper approximation. If $\underline{FC}(X) = \overline{FC}(X)$, then X is said to be exact with respect to covering approximation space(U,C). Otherwise, X is said to be covering rough set with respect to (U,C).

Definitions of lower and upper approximations of a set can now be easily formulated using tolerance classes. In order to do this, substitute tolerance classes for indiscernibility classes in the basic definition of lower and upper approximations of set. Thus, the tolerance approximations of a given subset X of the universe U is defined as in definition 4.

Definition 4: Let $X \subset U$ and a binary tolerance relation R is defined on U. The lower approximation of X, denoted by $\underline{R}(X)$ and the upper approximation of X denoted by $\overline{R}(X)$ are respectively defined as follows:

$$\underline{R}(X) = \left\{x \in X, R(x) \subseteq X\right\} \text{ and } \overline{R}(X) = \bigcup_{x \in X} R(x)$$

In this work, we propose an intelligent clustering algorithm using first type covering rough sets for clustering human activity sequence transactions for assisting robotics. Let $x_i \in U$ be a transaction consisting of sequence of human activity performed in various locations. For clustering user transactions, initially each transaction is taken as a single cluster. Let the i[th] cluster be $C_i = \{x_i\}$. Clearly, C_i is a subset of U. The covering based upper approximation of C_i , denoted as $\overline{CR}(C_i)$, is a set of human activity transactions similar to x_i . For any non-negative threshold value $\delta \in [0,1]$ and for any two objects $x, y \in U$, a binary relation τ on U denoted as $x \; \tau \; y$ is defined by $x \; \tau \; y$ iff $Sim(x, y) \geq \delta$,where Sim represents a similarity between x and y.This relation R is a tolerance relation, and R is both reflexive and symmetric but transitivity may not always hold.The first upper approximation $\overline{\mathrm{CR}}(\mathrm{x}_i)$ is a set of objects that are most similar to x_i. Thus, first upper approximation of an object x_i can be defined as follows:

Definition 5: For a given non-negative threshold value $\delta \in [0,1]$ and a set $\mathrm{X} = \{\mathrm{x}_1, \mathrm{x}_2,,,, \mathrm{x}_n\}$, $X \subseteq U$ the first upper approximation is

$$\overline{\mathrm{CR}}(\{\mathrm{x}_i\}) = \{\mathrm{x}_j \mid \mathrm{Sim}(\mathrm{x}_i, \mathrm{x}_j) \geq \delta\}$$

Some sets in the collection resulting from the first upper approximation may share elements (the so-called boundary elements). The boundary elements can guide the clustering process. The shared elements, generated after first upper approximation, may be the potential candidate of the new collection formed in the second or higher upper approximations. This can be decided by calculating the strength of shared element to all the clusters it belongs. This is measured using a parameter called relative similarity.

The value of the second and the higher similarity upper approximations is computed under the condition of relative similarity. For two intersecting sets $X, Y \in U$ X, The relative similarity of X with respect to Y is given by

$$\mathrm{Re\,lSim}(\mathrm{x}_i, \mathrm{x}_j) = \frac{\left|\overline{\mathrm{CR}}(\mathrm{x}_i) \cap \overline{\mathrm{CR}}(\mathrm{x}_j)\right|}{\left|\overline{\mathrm{CR}}(\mathrm{x}_i) - \overline{\mathrm{CR}}(\mathrm{x}_j)\right|} \text{ Where } \overline{CR}(x_i) \not\subset \overline{CR}(x_j)$$

Here the above relative similarity formula is rewritten based on first type covering based rough set.

Let x_i,x_j are two target set and, which are the subset of U, and Cover is C, as per the definition type 3, first covering upper approximation $\overline{\mathrm{FC}}(\mathrm{X})$ is used to calculate the relative similarity. The first type covering based relative similarity is represented as follows:

$$\mathrm{Re}\,lSim(x_i, x_j) = \frac{\left|\overline{FC}(x_i) \cap \overline{FC}(x_j)\right|}{\left|\overline{FC}(x_i) - \overline{FC}(x_j)\right|} \text{ Where } \overline{\mathrm{FC}}(\mathrm{x}_i) \not\subset \overline{\mathrm{FC}}(\mathrm{x}_j)$$

The relative similarity defined above, measures the ratio of size of the shared boundary between two sets and the number of elements that exclusively belong to the set under consideration. Now we define the proposed constrained-similarity upper approximation in the following definition:

Definition 6: Let $X = \{x_1, x_2,,,, x_n\}$, $\mathrm{X} \subseteq \mathrm{U}$. For a fixed non-negative value $\sigma \in [0,1]$, the constrained similarity upper approximation of x_i is given by

$$\overline{CRR}(\{x_i\}) = \left\{ x_j \in \bigcup_{x_l \in \overline{CR}(x_i)} \overline{CR}(x_l) \mid \mathrm{Re\,1}Sim(x_i, x_j) \geq \sigma \right\} \text{ Where } \overline{\mathrm{CR}}(\mathrm{x}_i) \not\subset \overline{\mathrm{CR}}(\mathrm{x}_j).$$

In other words, all the sequences x_i which belong to the similarity upper approximations of elements of $\overline{CR}(x_i)$ that are relatively similar to x_i are constrained (or merged) into the next similarity upper approximation of x_i .

The proposed covering rough clustering algorithm using refined first type covering based rough set is given in Algorithm 1.

The process is repted for computing successive constrained-similarity upper approximations for a given σ until two consecutive constrained-similarity upper approximations remain the same. Here, σ is a user-defined parameter called relative similarity, used to merge two upper approximations for the formation of the second and higher upper approximations. δ is computed dynamically by taking average of all the transaction's similarity values. It has been used to define the similarity between two objects and is utilized to find the first upper approximation. The constrained-similarity upper approximation is computed for all transactions of U. The running time complexity of the proposed algorithm is O(n^2).

5. EXPERIMENT AND RESULT ANALYSIS

In this paper, covering rough clustering approach is applied to recognition of unstructured human activity sequences for assisting robotics. Cornell activity data set CAD 120(Jaeyong Sung et al,2014) 8,20] is used for testing the proposed approach. CAD-120 data sets comprise of RGB-D video

Algorithm 1. Covering Rough Clustering algorithm

Input:
T: A set of n transactions $\in U \quad S_i \in US$
Threshold is computed dynamically based on the transaction mean
Relative similarity $\sigma \in (0,1)$
Output:
Cluster scheme C
Begin
Step 1: *Construct the similarity Matrix using* S^3M *measures.*
Step 2: *For each* $x_i \in U$, *Compute*
$S_i = \overline{FCR}(x_i)$ *using first type covering based rough upper approximation as per definition 2 for given computed Threshold* δ
Step 3: *Let* $US = \bigcup_i S_i$,
$C = \phi$
Step 4: For all Compute the Next constrained similarity upper *Approximation* S' *using Definition 3*
For similarity σ
If $S_i = S_i'$
$C = C \bigcup S_i'$
$US = US \setminus \{S_i\}$
Endif
Step 5: *Repeat step 4 until* $US \neq \phi$
Step 6: Return C
End

sequences of humans performing activities (Hema Koppula & Ashutosh Saxena,2015). Which are recording using the Microsoft Kinect sensor. The following 10 high-level activities are considered as sequence transactions.

…making cereal, taking medicine, stacking objects, unstacking objects, microwaving food, picking objects, cleaning objects, taking food, arranging objects, having a meal. Each sequence in the dataset corresponds to human activity performed in unstructured environment.

To illustrate our proposed approach, we consider the following data: S_1=' making cereal', S_2=' taking medicine ', S_3=' stacking objects', S_4=' unstacking objects', S_5=' microwaving food ', S_6=' picking objects ', S_7=' cleaning objects ', S_8 =' taking food ', S_9=' arranging objects ', S_{10}=' having a meal '. Considered the following 10 data sequences:

T1=(S_6 S_7 S_8 S_7 S_8 S_7),
T2=(S_2 S_3 S_4 S_7 S_9),
T3=(S_7 S_9 S_8 S_7 S_9 S_8),
T4=(S_1 S_2 $S_1$$S_2$ S_2 S_4),
T5=(S_6 S_5 S_8 S_8 S_8 S_9),
T6=(S_6 S_1 S_1 S_6 S_3 S_7),
T7=(S_1 S_3 S_4 S_1 S_2 S_1),
T8=(S_1 $S_3$$S_2$$S_1$ S_2 S_1 S_{10}),
T9=(S_2 S_1 S_6 S_7 S_8 S_{10})
T10=(S_1 S_2 S_6 S_7 S_8 S_{10})

Let us consider the universe U and cover C is as follows

$$U = \left\{T1, T2, T3, T4, T5, T6, T7, T8, T9, T10\right\}$$

$$\mathrm{C} = \left\{ \begin{Bmatrix} \mathrm{T1, T2, T3} \},\{\mathrm{T2, T3, T4,} \},\{ \mathrm{T5, T6} \} \\ ,\{\mathrm{T6, T7, T8, T9} \},\{\mathrm{T8, T9, T10} \end{Bmatrix} \right\}$$

The similarity table is computed using S^3M similarity metric with p=0.5 as shown in table 1
With the help of S^3M similarity table1, The first similarity upper approximation is computed using definition 5 and threshold value δ is given by $\overline{CR}\left(T_i\right)$ for i=1,2,…10. Here the threshold value δ is determined dynamically by calculating the average of all the transaction values:

$$\delta = \frac{1}{n}\sum_{i=1}^{n} T_i$$

The threshold value δ for transactions T1=0.15 which is calculated using above formula. The transactions which satisfies δ are included in the first similarity upper approximation. Such way all the transaction threshold values are calculated and first similarity upper approximation is computed. The value of the second and the higher similarity upper approximation is computed under the condition of relative similarity. The process is repted for computing successive constrained-similarity upper approximations for a given σ until two consecutive constrained-similarity upper approximations

Table 1. S^3M *Similarity metric computation for 10 transactions*

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
T1	1	0	0	0	.21	.30	0	0	0	0
T2	0	1	0	.47	.17	.17	.17	0	.16	.16
T3	0	0	1	0	0	.25	.33	.33	0	.21
T4	0	.47	0	1	.17	0	.45	.27	.24	.5
T5	.21	.17	0	.17	1	.18	0	0	0	0
T6	.29	.17	.25	0	.18	1	.18	.21	0	.17
T7	0	.17	.34	.45	0	.18	1	.58	.17	.62
T8	0	0	.34	.27	0	.21	.58	1	0	.5
T9	0	.16	0	.24	0	0	.17	0	1	.24
T10	0	.16	.21	.5	0	.17	.63	.5	.24	0

remain the same. Unchanging in the constrained-similarity upper approximations for all the element's coverge the algorithm. The final clusters are shown in Fig.1.

Experiments were performed on cornell activity data set and proposed covering rough clustering approach performance is good. Both, the number of overlapping clusters as well as the time taken to execute the program on various samples of randomly selected data sequences are observed.

Observation from Figure 2 shows that number of overlapping is increasing in covering rough clustering approach than exisiting clustering algorithm.

Figure 1. Covering based overlapping clusters

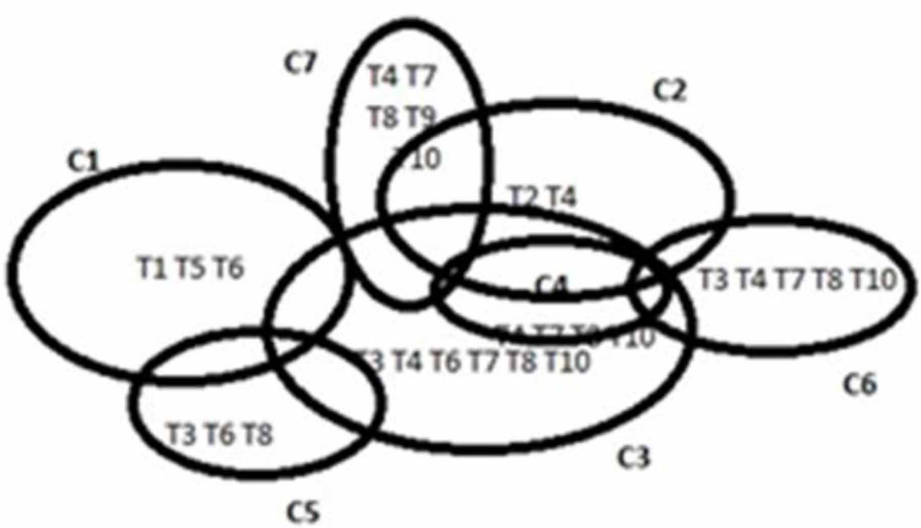
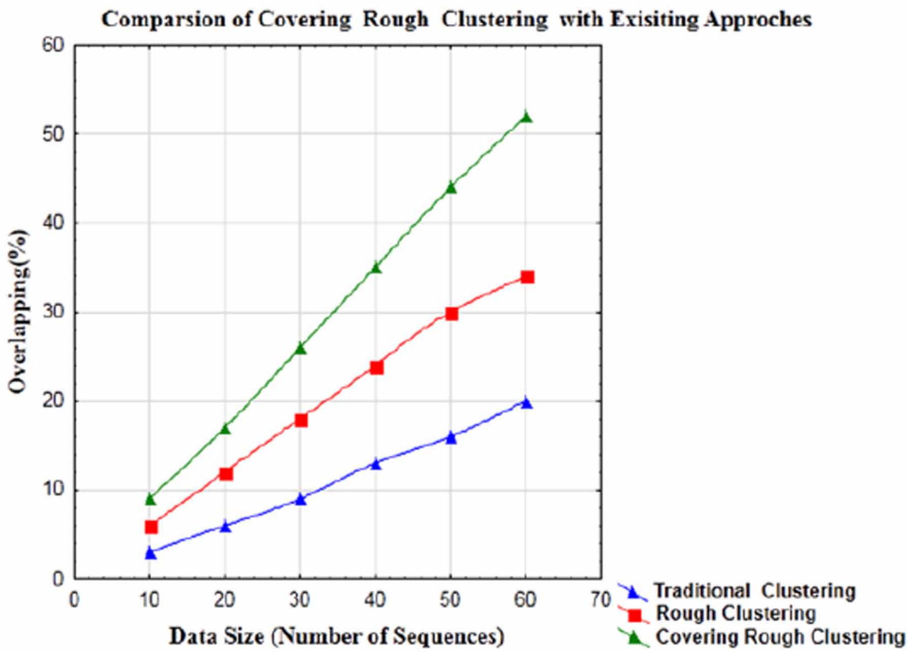


Figure 2. Performance of Covering Rough Clustering Approach



6. CONCLUSION

In this paper, covering rough clustering approach is applied to cluster human activity sequences using the concept of refined first type covering based rough sets. The experiments were performed on to recognition of unstructured human activity data sequences, which is used for assisting robotics. In this work, the threshold value δ is dynamically computed, based on this the first similarity upper approximation is computed. The concept of refined first type covering based is introduced to cluster sequential data. The obtained cluster will be useful to train the activities of the robot. The result shows that number of overlapping is increasing in covering rough clustering approach than conventional rough clustering approaches.

REFERENCES

Bergroth, L., Hakonen, H., & Raita, T. (2000). A survey of longest common subsequence algorithm. *Proceedings of theSeventh International Symposium on String Processing and Information Retrieval*, Atlanta, GA, USA (pp. 39–48). doi:10.1109/SPIRE.2000.878178

Bozkir, S., & Sezer, E. A. (2013). FUAT – A fuzzy clustering analysis tool. *Expert Systems with Applications*, *40*(3), 842–849. doi:10.1016/j.eswa.2012.05.038

Chatzis, S. P. (2011). A fuzzy c-means-type algorithm for clustering of data with mixed numeric and categorical attributes employing a probabilistic dissimilarity functional. *Expert Systems with Applications*, *38*(7), 8684–8689. doi:10.1016/j.eswa.2011.01.074

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys*, *31*(3), 264–323. doi:10.1145/331499.331504

Koppula, H., & Saxena, A. (2016). Anticipating Human Activities using Object Affordances for Reactive Robotic Response. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *38*(1), 14–29. PMID:26656575

Krishnasamy, G., Anand, J. K., & Raveendran, P. (2014). A hybrid approach for data clustering based on modified cohort intelligence and K-means. *Expert Systems with Applications*, *41*(13), 6009–6016. doi:10.1016/j.eswa.2014.03.021

Li, F., Ye, M., & Chen, X. (2014). An extension to Rough c-means clustering based on decision-theoretic Rough Sets model. *International Journal of Approximate Reasoning*, *55*(1), 116–129. doi:10.1016/j.ijar.2013.05.005

Lingras, P., & West, C. (2004). Interval set clustering of web users with rough k-means. *Journal of Intelligent Information Systems*, *23*(1), 5–16. doi:10.1023/B:JIIS.0000029668.88665.1a

Mitra, A., Satapathy, S. R., & Paul, S. (2013).Clustering analysis in social network using covering based rough Set. *Proceedings of theIEEE International Advance Computing Conference* Ghaziabad, India (pp. 476-481). doi:10.1109/IAdCC.2013.6514272

Pawlak, Z. (1982). Rough Sets. *International Journal of Computer & Information Sciences*, *11*(5), 341–356. doi:10.1007/BF01001956

Pawlak, Z. (1991). *Rough Sets: Theoretical Aspects of Reasoning about Data*. MA, USA: Kluwer Academic PublishersNorwell. doi:10.1007/978-94-011-3534-4

Peters, G., Crespo, F., Lingras, P., & Weber, R. (2013). Soft clustering – Fuzzy and rough approaches and their extensions and derivatives. *International Journal of Approximate Reasoning*, *54*(2), 307–322. doi:10.1016/j.ijar.2012.10.003

Pradeep Kumar, P. Radha Krishna, Raju., S. Bapi and Supriya Kumar De. (2007). Rough clustering of sequential data. *Data & Knowledge Engineering*, 63(2), 183-199.

Saneifar, H., Bringay, S., Laurent, A., & Teisseire, M. (2008).S2MP: Similarity Measure for Sequential Patterns. *Proceedings of the Australasian Data Mining Conference*, Australia (*Vol. 87*, pp. 95-104).

Shi, Z., & Gong, Z. (2010). The further investigation of covering-based rough sets: Uncertainty characterization, similarity measure and generalized models. *Information Sciences*, *180*(19), 3745–3763. doi:10.1016/j.ins.2010.06.020

Sung, J., Selmanand, B., & Saxena, A. (2014). Synthesizing Manipulation Sequences for Under-Specified Tasks using Unrolled Markov Random Fields. *Proceedings of theIEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, Illinois, USA (pp. 2970-2977).

Tripathy, B. K., & Mitra, A. (2009).Some topological properties of covering based rough sets. *Proceedings of the IEEE – International conf. on Emerging Trends in Computing*, Virudhnagar, India (pp. 263-268).

Tripathy, B. K., Mitra, A., & Ojha, J. (2008). On Rough Equalities and Rough Equivalences of Sets. In *Rough sets and current trends in computing, LNCS* (Vol. *5306*, 92–102). *Springer-Verlag Berlin Heidelberg*.

Tripathy, B. K., & Tripathy, H. K. (2009). Covering Based Rough Equivalence of Sets and Comparison of Knowledge. *International Association of Computer Science and Information Technology*, *303–307*, 17–20.

Yao, Y., & Yao, B. (2012). Covering based rough set approximations. *Information Sciences*, *200*, 91–107. doi:10.1016/j.ins.2012.02.065

Zakowski, W. (1983). Approximation in space(U, Π). *Demonstratio Mathematica.*, *16*, 761–769.

Zhang, L., Pedrycz, W., Lu, W., Liu, X., & Zhang, L. (2014). An interval weighed fuzzy c-means clustering by genetically guided alternating optimization. *Expert Systems with Applications, 41*(13), 5960–5971. doi:10.1016/j.eswa.2014.03.042

Zhu, W., & Wang, F.-Y. (2007). On Three Types of Covering-Based Rough Sets. *IEEE Transactions on Knowledge and Data Engineering*, *19*(8), 1131–1144. doi:10.1109/TKDE.2007.1044