

An Efficient Trajectory representative generation moving data prediction using different Clustering *algorithm*

Vishnu Kumar Mishra

JNTU University, Hyderabad

Megha Mishra

CSVТУ University, Bhilai

Bhupesh Kumar Dewangan (✉ b.dewangan@ddn.upes.ac.in)

Department of Informatics, School of Computer Science, University of Petroleum and Energy Studies, Energy Acres, Dehradun, Uttarakhand

Tanupriya Choudhury

Department of Informatics, School of Computer Science, University of Petroleum and Energy Studies, Energy Acres, Dehradun, Uttarakhand

Research Article

Keywords: Trajectory clustering, Density based cluster, Framework grouped, Partition and group work, Data mining

Posted Date: April 19th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-439136/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

This paper highlighted Moving and Trajectory Object Cluster (MOTRACCLUS) algorithm and analysis the sub-trajectories and real-trajectories algorithm for moving data and suggest a new approach of moving elements. This paper evaluates the Hurricane data measure and mass less data measure entropy of trajectories objects of moving data of Chhattisgarh location. The paper covered Prediction Generation with their distance cluster *minimum description length* (MDL) algorithm and others corresponding distance cluster (CLSTR) algorithm. This paper highlighted the k-nearest algorithm with least cluster section (LCSS) model and dimensional Euclidean of MDL algorithm. Our algorithm consists of two parts that is partitioning and grouping phase. This paper develops and enhances a cluster of trajectory object and calculates the actual distance of moving object. This algorithm works on the CLSTR algorithm and calculates Trajectory movement of object. In this we evaluate the entropy of moving object by consideration the heuristic parameter.

1. Introduction

This papers we inspect the problem of discover similar trajectories of moving objects. The trajectory of a moving object is typically modeled as a sequence of consecutive locations in a multidimensional (generally two or three dimensional) Euclidean space. We are facing an unprecedented proliferation of mobile devices, many equipped with positional technologies such as Global Positioning System (GPS).

Moreover, the accessibility as well as the wide-ranging collection of applications these knowledge apparatus, has maintained their inclusive dispersion of everyday life, facilitate the frequent aggregation of moving objects data into large wide-ranging datasets. These promising datasets are stressing the implication of supervision and organize cluster of representative data. Technique used for remove information from database or cluster set data is called trajectory data mining.

The current developments in peripatetic database need the different electronic device like GPS system, Cellular mobile system and sensor technology and different tools for awareness, also perform very vital role for data analysis for performing various task related to data. For real life example, the mining computing the users equipped with movable devices that move in everywhere space and provide the record, which registered with device location at different time moments via some movable temporal databases.

2. Related Work

In this Section, we study previous work based on two major discovery techniques, Markov chain models and spatiotemporal data mining, for extracting movement patterns of an object from historical trajectories [1-3]. Markov chain models have been widely used in order to estimate the probability of an object's movements from one region or state to another at next time period. Ishikawa et al. derive the Markov transition probabilities between cells from indexed trajectories [4, 5].

In their further study [6, 7], a special type of histogram, called mobility histogram, is used to describe mobility statistics based on the Markov chain model. They also represent the histogram as cube-like logical structures and support an online analytical processing (OLAP)-style analysis. Authors in [8] classify an object's mobility patterns into three states (stationary state, linear movement, and random movement) and apply Markov transition probabilities to explain a movement changes one state to another consider the location tracking problem in PCS networks[9-11]. Both studies are based on the same Markov process in order to describe users' movements from one or multiple Personal Communications Services (PCS) cells to another cell. However, they have different ways to model users' motilities using Morkov models, thus, show distinct results to each other [12, 13].

Data mining play important role to discover the knowledge and some time it is also called knowledge discovery, some time this technique help to extracting the data from large database [14, 15]. Data mining clustering approach generate answer and classifying movement analysis of data by using some clustering and prediction algorithm [16, 17].

Data mining only a tool and framework which help to mine temporal and pictorials trajectory moving data. Through the classified trajectory moving data can be represent by cubes and relational transaction approach [18-20]. The survey of instance of data mining methods based on classical relational and transactional data can be found in traditional clustering trajectory movement [21, 22] and [23]. Based on that trajectory mining cluster many researchers suggest the many technologies Global Navigation Satellite System (GNSS) and Radio frequency identification (RFID) [24, 25].

The trajectory data mining generation is analysis tasks based, suggest visualization of Two-dimension (2D) and Three-dimension (3D) data movement. The trajectory mining moving data classify three domains: these are operational, dynamic and social trajectory motion of data [26, 27]. This survey represents different algorithm efficient clustering generation with respective density based clustering algorithm [28, 29].

The trajectory data and Clustering are the generally measured as the field of data mining. In data mining classification and Clustering play main role. The Clustering algorithms can be divided into four different types [30, 31]: The first method generally based on density monitoring method that real life example Density-based spatial clustering of applications with noise (DBSCAN) and other example is Ordering points to identify the clustering structure (OPTICS) algorithm. The second type is grid-based algorithm; the example of this algorithm is Statistical Information Grid (STING) [32-34]. The third algorithm *k*-means which is example of partitioning algorithm and fourth is hierarchical algorithm, one of the best examples of hierarchical algorithm is Balanced iterative reducing and clustering using hierarchies (BIRCH) algorithm. Suggested algorithm Trajectory clustering (TRACCLUS) follow the rule group of density-based methods[35-37].

The fig 1. Shows probability of trajectory data movement of different spot with respective time span. The different spot colors represent actual predictive generation of object. The spot black and brown represent the maximum data movement of object and center of attention the same cluster generation. The dataset DBM based clusters algorithm is provinces of high density disjointed by provinces of low-density algorithm that is example of *k*-means [38, 39]. It is also observed that density-based clustering algorithm a part of DBSCAN algorithm. As mainstream preceding investigation has take up be partitioning and group of apex statistics. Trajectory representative generation moving data prediction has maximum and given projected data clustering suggested by Gaffney *et al.* [40, 41]

The fig 2. correspond the overall similar trajectory data object movement at density based clustering at per 500 kilometer range with respect to latitude and longitude . This latitude and longitude help for actual movement of object in each direction with classify distance 500 per kilometer. This emphasis the modification of the representative generation moving probabilistic modeling algorithm monthly basis i.e. is called TRACCLUS algorithm. Properly, the probability calculation based on representative generation prediction.

3. Algorithm Criteria

3.1. The *CLSTR* Algorithm

The TRACCLUS trajectory clustering algorithm contains the three different phases. These are trajectory planning (line 2), Line Segments Clustering (line 4) and Trajectory generations. These phases execute in sequentially with minimum trajectory movement i.e. the first part we calculate the partitioning and then second phase we execute the trajectory clustering. We comprehensive clarify these algorithms below.

Algorithm Generation *CLSTR* (Clustering Trajectory)

Input: Find trajectories generation data $T = \{G_1, \dots, G_{nm|n}\}$

- Output: (1) set prediction data $R = \{S_1, \dots, S_{n_{clp}}\}$
 (2) A $S = \{\text{set of trajectories moving set of data}\}$

Algorithm:

/* CLSTR clustering prediction */

- 01: $\forall (G \in T)$ recursive until termination ;
 02: *Predictive generation Trajectory clustering;*
 find L_s and $N\epsilon(L)$;
 03: Accumulate L_s into a set E ;
 /* Grouping Phase */
 04: Execute *Line Segment Clustering for E;*
 Find value of R clusters ;
 05: search belong ($S \in R$) do;
 06: Repeat *Predictive generation Trajectory clustering;*

3.2 Partitioning Trajectory (Partitioning Phase)

This portion covers the *Approximate Trajectory Partitioning* algorithm for trajectory clustering for data. The algorithm *Approximate Trajectory Partitioning* shows below. This work computed $MDL_{no\text{par}}$ and MDL_{par} for respectively argument in a trajectory. Compare actual value of each MDL_{par} and $MDL_{no\text{par}}$ out of them find out maximum value then put in the instantly value to $MDL_{\text{par}}(\text{pmdlstaIndex}, \text{pmdlcurrentIndex})$. Set the current appxomately value in stating and repeat CP_i as startmdlP1.

Algorithm Prediction generation Trajectory MDL

- Input: A prediction generation $PG_i = g_1g_2 \dots g_n \dots ppgin_i$
 Output: Prediction generate upto last $C_p = \{C_1, c_2, \dots, C_{ppgin_i}\}$

Algorithm:

- /* stating : Prediciton Generation */
 01: Initialize the value CP_i in srt P_1, p_2, \dots ;
 02: set $Indexvaluestr := 1$, $maximumlength := MDL_{\text{par}}$ $mimum = P_{curr} + currIndex$;
 03: for ($maxIndex$ start+ $maximunvalue \leq currentIndxggu$); repeat unitl termination
 04: $maxcuurrvalue = pccggstart + current$ start;
 05: $max_{\text{par}} := CDMDL_{\text{total}}_{\text{par}}(\text{maxcuur srart}, \text{max curr index})$;

```

06: minnopar := CDMDLtotalnopar (maxcuur sart, max curr index);
/*Find max and min patterned the separating existing argument that generation as CDMDL totalnonpa */
07: CDMDL totalpar= pccggstart -1
/* Prediction generation of previous point */
08: Average= pccggstart -1;fing generation each;
09: pccggstart := pcccurrindexst-1, maxlength :=1;
10: recursive
11: total max := pccggstart-1;
12: Assign the value pleni into Cp /* the last */

```

3.3 Clustering Trajectory (Grouping Phase)

This portion, we recommend a different line algorithm, especially emphasis on line *segment clustering algorithm* for moving data trajectory clustering. We also focus on trajectory density prediction generation algorithm for trajectory clustering. Suppose there are segmentation of line belong to S and produces a set of clusters that is O. For defining the cluster set parameter we need *index value cluster Id* two data set i.e. is $S\epsilon(L)$ and Lnr_1, \dots, L_{nmcls} . These approaches suggest clustering trajectory algorithm provides different appearances of algorithm DBSCAN with many parts in trajectory clustering.

Algorithm: K-nearest prediction generation density Clustering

Input: (1) K-nearest Line prediction generation segments $D = \{G_1, \dots, G_{nmcls}\}$,
(2) set K-nearest prediction $S\epsilon(L)$ and Lnr_1, \dots, L_{nmcls} .

Output: K-nearest Line prediction generation segments $S = \{s_1, \dots, s_{nmcls}\}$.

Algorithm:

```

/* prediction starting point generation */
01: Assign generation cluster Id to be -1;
02: Identify all unbreakable cline and set cluster Id value;
03: iterative belong ( $L \in D$ ) until
04: go (S each unbroken Id) jump
05: assign  $S\epsilon(D)$ ;
06: put ( $|S\epsilon(D)| \geq GNmax$ ) repeat
07: provide sum of  $rS \{A\} \in Sr(D)$  on Id;
08: Addition  $N\epsilon(I) - \{A\}$ ;
/* 2 Step prediction generation */

```

```

09:    assign cluster Id (S, cluster Id, r, MLmax);
10:    put cluster Id as GNmax; /* generate new id */
11:    do;
12:    drop unstasidied generation(remark);

/* 3 Step prediction generation */
13: Assign r  $G_s \in D, S_{clusterid}$ ;

/* Cluster predication trajectory */
14: if ( $D \in G$ ) till

/* from stating check clusterid*/
15: for ( $|KCN(G)| > MLmax$ )
16:    check unsatisfied value and remove from cluder id;
17: Increase value(S, D, G, MLins) {
18: for each ( $D=g$ ) till
19:    each pass value assign in Sg;
20:    add  $G \alpha (S)$ ;
21:    for  $\{A\} \in Sr(D)$  on Id) do
22:        ITERATION (  $NS\alpha (g \setminus D)$  reperat
23:        set (G is actual cluster value) print
24:        value of cluster Id to G;
25:        set (S is removed cluster id) print
26:        value of cluster Id to D;
27:    Stop removing value from G to value  $p_1, p_2, \dots, p_n$ ;
28:    Stop prediction generation from S;
29: }

```

Algorithm: DBSCAN Trajectory partitioning clustering

Input: (1) Create new cluster *generation predication* P_i

(2) Find *MLins* value for cluster *generation*

(3) Set prediction φ .

Output: Set demonstrative trajectory T_i prediction generation for maximum P_i .

Algorithm:

- 01: Find max value of direction vector field ;
- 02: Alternate the hatchets X axis is equivalent to;
- 03: Now set starting and ending value for cluster P_i
- /* coordinate of the c' axis donation */
- 04: average value G with new generation p' -values;
- 05: if ($T \in G$) recursive
- /* calculate all cluster prediction $_p$ */
- 06: sum mov_p segments that contain the p' value of the predication r ;
- 07: if ($mov_p \geq pLinuus$) then
- 08: rel in p' values between t and its immediately previous point f ;
- 09: put ($max \geq r$) get
- 10: total max generation $movavgC_p$;
- 11: get max and min value mov_p ;
- 12: Append avg_p to the end of RTR_i ;

3.4 Formalization Using the MDL Principle

This area suggests a formalization trajectory prediction of optimum trade between precision versus terseness. This portion also emphasis the accept the minimum prediction generation description density length (MDL) standard extensively.

The MDL classify into two different parts, these are $G(P)$ and $L(G|T)$ where T emphasis trajectory movement and G belongs maximum prediction data. There binary apparatuses are casually stated as follows [if ($diff \geq \alpha$)]: we formulate $G(P)$ by Formula (1). Here, $len(p_{cj+1})$ denotes the length of a line segment $pcj+1$, i.e., the Euclidean distance between pcj and $pcj+1$. Hence, $L(H)$ signifies the amount of the distance of all trajectory partitions.

$$L(H) = \sum_{j=1}^{par i - 1} \log_2(len(p_{cj}p_{cj+1})) \quad (1)$$

Approximate Solution

The algorithm Approximate Trajectory Partitioning shows below. Here we calculate GTR for cluster prediction of moving data. Here value of $CLCSTR_{par}$ is assign maximum trajectory area i.e. belong $\{A\} \in Sr(D)$ on Id_{nopar} , assign actual data with its adjacent value of moving data , for every point $\{A\} \in Sr(D)$ on Id . Then, we recurrence generation data practice for

approximate solution i.e. start with initial point (startInex: = currIndx-1, length: = 1). Representative $CLCSTR_{par\ present}$ provides best generation approximation clustering.

Algorithm: New Representative $CLCSTR_{par\ present}$ Generation

Input: (1) Consider PLT_i as $CLCSTR_{par\ present} = \{c1, c2 \dots Cnindexstart\}$;

(2) $MLins$ (3) A smoothing parameter α

Output: The demonstrative PLR_i as $CLCSTR_{par\ present} = \{p1, p2 \dots pnindexstart\}$;

Algorithm:

01: Find max value of direction vector field ;

02: Replace the hatchets X axis is equivalent to ;

03: set value $CLCSTR_{par\ present} = currentindex$;

04: iterative belong ($L \in D$) until;

05: if ($G \in r$) repeat

06: Let num_p be the number of the line segments that contain the X -value of the point p ;

07: if ($num_p \geq G\ MLins$) then

08: $calu$ difference in X -values between p and its nearest point;

09: if ($diff \geq \alpha$) then

10: Compute the average coordinate avg_p ;

11: Unwrap the spin and find out avg_p ;

12: Append avg_p to the end of RTR_i ;

Clustering Neighbors K-Clustering (NK-CN)

The set od points (p , int k , real num α)

// where α is greater than zero ($\alpha > 0$).

BEGIN

$p =$ Find out core value P (int p , k , α);

if $pointp \neq null$ then

$Cul\ 1\ set = Set1$ (int p , k , α) set at initial point;

$Clu1 = Set$ and $GetCluID1$ ();

$C1 = Set$ and $GetInitial\ valueClu1$ (for p , tp , $Set1$, k , $cluID1$);

```

    Release Cluster1 (p, C1 , Core1, α,K);

Start clustering;

D =D ∪ nk-distance (point) (point);

CoreSet1 = Coreset1 (object); Release clur1 (Set p, C1,Set1, int k, float α)

BEGIN

    Seedset1 = CoreSet1;

    while not SeedSet1.empty() DO

        Pointp = GetOutPointp1 (SeedSet1);

            CoreSet1(object1) = D

                SeedSet1 = SeedSet1 ∪ (object);

                CoreSet1 = Coreset1 ∪ (object);

            terminating;

        close each iteration;

    D =D ∪ nk-distance (point) (point);

    terminate loop;

end Expandclus1.

```

The *TRCLS* trajectory clustering algorithm contains the three different phases. These are trajectory planning (line 2), Line Segments Clustering (line 4) and Trajectory generations. These phases execute in sequentially with minimum trajectory movement i.e. the first part we calculate the partitioning and then second phase we execute the trajectory clustering. We comprehensive clarify these algorithms below.

Algorithm Prediction generation Trajectory MDL

Input: A prediction generation $PG_i = g_1g_2 \dots g_n \dots ppgin_i$

Output: Prediction generate upto last $C_p = \{ C_1, c_2, \dots, C_{ppgin_i} \}$

Algorithm:

/* stating : Prediciton Generation */

01: Initialize the value CP_i in srt P_1, p_2, \dots ;

02: set $Indexvaluestr := 1$, $maximumlength := MDL_{par}$ $mimumin = P_{curr} + currIndex$;

03: for ($maxIndex$ start+ $maximunvalue \leq currentIdxggu$); repeat until termination

04: $maxcuurrvalue = pccgggstart + current start$;

```

05:   maxpar := CDMDL totalpar(maxcuur sstart, max curr index);
06:   minnonpar := CDMDLtotalnonpar (maxcuur sstart, max curr index);

/*Find max and min patterned the separating existing argument that generation as CDMDL totalnonpa */

07:   CDMDL totalpar= pccgggstart -1

/* Prediction generation of previous point */

08:   Average= pccgggstart -1;fing generation each;
09:   pccgggstart := pcccurrIndexst-1, maxlength :=1;
10:   recursive
11:   total max := pccgggstart-1;
12:   Assign the value pleni into Cp /* the last */

```

Algorithm: DBSCAN Trajectory partitioning clustering

Input: (1) Create new cluster *generation predication* P_i

(2) Find MLins value for cluster *generation*

(3) Set prediction φ .

Output: Set demonstrative trajectory T_i prediction generation for maximum P_i .

Algorithm:

```

01: Find max value of direction vector field ;
02: Alternate the hatchets  $X$  axis is equivalent to;
03: Now set starting and ending value for cluster  $P_i$ 

/* coordinate of the  $c'$  axis donation */

04: average value G with new generation  $p'$ -values;
05:   if ( $T \in G$ ) recursive

/* calculate all cluster predictionp */

06: sum  $mov_p$  segments that contain the  $p'$  value of the predication  $r$ ;

07:   if ( $mov_p \geq pLincuus$ ) then
08:      $rel$  in  $p'$  values between  $t$  and its immediately previous point  $f$ ;
09:     put ( $max \geq r$ ) get
10:   total max generation  $movavgC_p$ ;

```

11: give max and min value mov_{p_i} :

12: Append avg_p to the end of RTR_i

4. Experimental Evaluation

The experiment performed using the approximation algorithm. The main aim of our research work is to compare the result which we obtain by using CLSTR algorithm and the previous traditional approaches. The following points are considerably important for performing experimental work

The distance vector selected for Euclidean value will always remain unchanged throughout the research work. The comparison is performed for each and every pair of sequence value taken and tried out to uncover the minimum possible gap between the paired values. The value accepted for the minimum distance vectors.

Table 1. Correspondence values of Prediction Generation with Probability

Distance Vector (Km)	Prediction Generation /probability	ALGORITHM					
		MDL	SCAN	DBSCAN	CLCSTR _{par} present	NK-CN	K-nearest
500	4.05/0.08	0.42	0.736	0.0043	0.00490	23.905	0.454
1000	6.01/.14	0.506	0.673	0.00341	0.0042	18.907	0.876
1500	5.04/0.06	0.358	0.631	0.0026	0.0042	45.452	0.0048
2000	5.08/0.12	0.323	0.588	0.00271	0.00271	42.454	0.451
2500	4.02/0.15	0.617	0.453	0.0032	0.0032	94.256	0.349
3000	4.06/0.03	0.806	0.5935	0.0043	0.0043	93.245	8.651
3500	5.01/0.54	0.7	0.7	0.323	0.588	0.7	0.67

Table 2. Correspondence values of distance cluster MDLdis

Distance cluster						
Comapre	Actual time	Representative generation				
		5	15	20	26	
MDLdis						
5	0.45	0.42	0.42	0.736	0.0043	0.00490
5	0.6	0.876	0.506	0.673	0.00341	0.0042
3	0.4	0.631	0.358	0.631	0.0026	0.0042
3	0.6	0.588	0.323	0.588	0.00271	0.00271
6	0.45	0.42	0.736	0.0043	0.00490	0.0032
6	0.7	0.876	0.806	0.5935	0.0043	0.0043

Table 3. Correspondence distance cluster CLSTRalgo.

Cluster generationtime cycle in (sec)						
Comapre	Actual time		Representavie generation			
CLSTR algo			5	10	26	52
4	0.67	0.67	0.358	0.631	0.67	0.0043
4	0.7	0.876	0.806	0.5935	0.0043	0.0043
3	0.4	0.631	0.5935	0.0043	0.0043	0.7
3	0.6	0.00271	0.00271	0.588	0.00271	0.00271
5	0.45	0.00490	0.0032	0.0043	0.00490	0.0032
5	0.7	0.67	0.806	0.5935	0.0043	0.0043

Next step we applied for DTW and applied the CLSTR algorithm for fin ding the more feasible values of co-ordinates. Distance vector is calculated for above mentioned two traditional approaches but it seems to take the optimized value of k means vector.The other hand in CLSTR algorithm no needs to manually optimize the parameter it automatically select the best suited one. The reason behind this is that it compares each and every possible sequence of pairs and there distances also best algorithm.

Table 4. Correspondence distance best algorithm for values of Prediction Generation with Probability

Distance Vector (Km)	Prediction Generation /probability	ALGORITHM(distance density clustering)					
		MDL	SCAN	DBSCAN	CLCSTR _{par} present	NK-CN	K-nearest
500	4.05/0.08	yes	yes	yes	yes	yes	no
1000	6.01/.14	yes	yes	no	yes	no	no
1500	5.04/0.06	yes	yes	yes	yes	yes	no
2000	5.08/0.12	yes	no	yes	yes	yes	no
2500	4.02/0.15	yes	no	yes	yes	yes	yes
3000	4.06/0.03	yes	no	yes	yes	yes	yes
3500	5.01/0.54	yes	yes	no	yes	no	no

The value selection of ω is very important for representation of LCSS .The experiment is approximately performed 15 times on an average to find out the most suitable clustering sequence with and without CLSTR algorithm and then the result is compared. CLSTR algorithm need not require any kind of sampling.

Table 5. Correspondence values and running times cycle between two sequences from our ANIMALS dataset

ω	φ	Correspondence					Running time cycle in (sec)				
		Actual time	No of tries movement				Actual time	No of tries movement			
			5	10	26	52		5	10	26	52
4	0.43	0.319	0.42	0.736	0.0043	0.00490	23.905	0.454	0.617	0.453	0.0032
4	0.7	0.451	0.506	0.673	0.00341	0.0042	18.907	0.876	0.806	0.5935	0.0043
6	0.67	0.349	0.358	0.631	0.0026	0.0042	45.452	0.0048	0.0032	0.0032	0.00491
6	0.7	8.651	0.323	0.588	0.00271	0.00271	0.7	0.451	0.506	0.673	0.00341
7	0.45	0.454	0.617	0.453	0.0032	0.0032	94.256	0.349	0.358	0.631	0.0026
7	0.6	0.876	0.806	0.5935	0.0043	0.0043	93.245	8.651	0.323	0.588	0.00271
5	0.7	0.451	0.506	0.673	0.00341	0.7	0.451	0.506	0.673	0.00341	0.0042
5	0.45	0.454	0.358	0.471	0.00341	0.0042	0.358	0.631	0.5935	0.0043	0.0032

5. Result And Discussion

The main objective of this work is to carry out the analytical discussion of data and also watching the entropy particulars of it. The work emphasizes the capacity of object when it is on movable condition and its actual efficiency to carry out the k-measure capacity. The value of φ we obtain from above work we can see that is 0.43 which is the lowest one observed during the analysis part. The Time complexity of our MOTRACCLUS algorithm is $O(n)$ as liner. Prediction generation Trajectory MDL and K- nearest prediction generation density Clustering is $O(n \log n)$ - Linearithmic algorithm while DBSCAN Trajectory partitioning clustering and DBSCAN Trajectory partitioning clustering $O(n^2)$ – linear Quadratic time.

From the Fig 3. we can observe that the value which we have calculated for $\varphi = 41$ and $\omega = \text{aprox}(15)$ not showing much similarity and also represent the different aspect when we will increase the quality value of $\omega = \text{aprox}(15)$. also we can see that if we will be taking the very small values of k it will give the better result as compared to large volume values. The range which is in between the values of 37 and 43 represent the most optimistic result of $T \varphi$.

Fig 4. Shows massless data measures. The on an average value we got for above analysis is $\text{AVG}[\omega(\varphi)] = 7.26$ when we set the values as $\Omega = 0.43$ and $\omega = \text{aprox}(15)$. So from above calculation the calculated value we achieved using the best suited parameter values are $\varphi = 41$ which is the most significant value obtained during the analysis phase and also matching with the value which we want to carry out for this work.

Fig 5 and Fig 6 shows hurricane data measure entropy and massless data measure entropy. The graph also represents the clustering outcome value which is $\omega = 11$ and it is not possible for all k points that clusters value should be 10. so from the above calculation and observation we can say that it is the most optimistic value which we got from the calculation and it is always good enough and produce the most quality clustering output for always $\omega = 11$. It also the close observation from above graph and calculation that no other value of ω can produce such a result which we got from above analysis.

6. Conclusion

This research work provides a way of 41 variations which we have used during the research work to form a more novel and optimistic approach for clustering configuration. The work which is carried out using different parameters and moving towards the most optimistic approach is based on the algorithm which we have used here CLSTR. The algorithm CLSTR is the best one which is designed over here taking the different value of parameters and choosing the most optimistic one. The work carried out is started with the division segments of different lines and parameters and then grouping the parameters on

the basis of similar point values. The main objective and advantage of the designed algorithms CLSTR is to find out the similar points of clusters from different database for moving objects based on some common values. Two datasets used for carrying out the research work which is based on the analytical data sets and animal datasets which is also real life dataset and costly one. The research work which is carried out has the main focus to find out the values most suitable and provide the optimistic result. One more effort which we applied here is that to find out the actual distance between clusters to validate their occurrences. The graphical aspect also represents the output which is based on the grouping of clusters that we have already stated in our algorithm CLSTR concluding the result which we got from the designed algorithm CLSTR we can say that this is the new approach towards the clustering paradigm and the objects which are movable. Data analytics which has been used is also insight the new approach towards the analysis of data. During the working period of this algorithm we found many issues and some other parameters for which we are working till now taking as a challenge.

Declarations

Acknowledgement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Zhegown (2020). Trajectory data mining of moving data. Inderscience publication ,Springer Intelligence System and Technology ,vol 10,no.4 , page 45.
2. Wang Xin,Ge Cui et.al , (2020) MDFI Sensors, A trajectory Collaboration Based Map Matching Approach for Low-Sampling –Rate GPS trajectories, 2020, 20-02057.
3. Wang,Y.Shen,Z.Bao,and X.Qin.Intelligent traffic analytics :from monitoring to controlling to controlling.In WSDM, pages 778-781, 2020.
4. A. Kuliko, Computer Systems That Classification and Prediction Methods from Statistics, Neural Networks, 2017.
5. Mehrotra and K. Chakrabarti , In Proc. of LDB 2018 ,Local dimensionality reduction: A new approach to indexing high dimensional spaces., pages 189–200.
6. Zheng (2020). Trajectory data mining overview. ACM Transactions on Intelligence System and Technology ,vol 6,no.3 article 29-51.
7. Yuan, S. Xia,L.Zhang,Y.Zhou, and C.ji, (2019). An efficient trajectory-clustering algorithm based on an index tree. Transaction of the Institute of Measurement and Control , vol.34, no.7, pp.850-861 (2019)
8. Wang,Z.Bao,J.S.Culpepper,Z.Xie,Q.Liu,and X.Qin. Torch:A search engine for trajectory data.In SIGIR, pages 535-544,2018.
9. Wang, G. G., Gandomi, A. H., Yang, X. S., and Alavi, A. H. (2016)A new hybrid method based on krill herd and cuckoo search for global optimization tasks, International Journal of Bio-Inspired Computation, Vol 8(5), pp. 286-299, 2016.
10. Sozio,T-h.,H. Chan and A. Guerquin.(2018) Fully dynamic k-center clustering.InWWW,pages 579-587,2018.
11. Al-Sharif, Z. A., Jararweh, Y., Al-Dahoud, A., and Alawneh, L. M. ACCRS(2017): autonomic based cloud computing resource scaling, Cluster Computing, Vol 4(1), Vol 3, pp. 1-10, 2017.
12. Lander ES, Golub TR, Mesirov JP (2018) An analytical method formulti-class molecular cancerclassification. SIAM Rev 45:706–723.
13. Yang,C.S. Jensen L.Chen and L .Gao.(2017) .Pivot basedmertix indexing,PVLDB10(10):1058-10969.2017
14. Han, J. and Kamber, M., Data Mining: Concepts and Techniques, 2nd ed., Morgan Kaufmann, 2016.
15. A. Kuliko, Computer Systems That Classification and Prediction Methods from Statistics, Neural Networks, 2017.
16. Lander ES, Golub TR, Mesirov JP (2018) An analytical method formulti-class molecular cancer classification. SIAM Rev 45:706–723.

17. Wu, G., et al: LNCS, vol.10178, pp.477-493.Springer, Cham (2017): Supporting real-time analytic queries in big and fast data environment.
18. Gill, S. S., Chana, I., Singh, M., and Buyya, R. CHOPPER: an intelligent QoS-aware autonomic Resource management approach for cloud computing, Cluster Computing, pp. 1-39, 2017.
19. Wang, H., and Yi, J. H. An improved optimization method based on krill herd and artificial bee colony with information exchange, Memetic Computing, Vol 10(2), pp. 177-198, 2018.
20. K. Mishra, Bhupesh Sahu et al. , Journal Name: International Journal of Advanced Research in Computer and Communication Engineering, Clustering the Movement of Trajectory Based Data, Volume 6 Issue 11, November 2017.
21. K. Mishra, Bhupesh Sahu et al. , Journal Name: International Journal of Advanced Research in Computer and Communication Engineering, Analysis the Space Partitioning and Group Clustering, Volume 6 Issue 11, November 2017.
22. Giannotti, F., Pedreschi, D.: Springer, Heidelberg 2008, Mobility, Data Mining and Privacy: Geographic Knowledge Discovery pp.189-195 .
23. Mehrotra and K. Chakrabarti , In Proc. of LDB 2018 , An generation of object analytical method formulate-class molecular cancer classification., pages 89–95.
24. K. Mishra, Bhupesh Sahu ,. Journal Name: International Journal of advanced Research in Computer Science and Software Engineering. Conceptual Studies of Intelligent Character Recognition Identified by IC Algorithms, Volume 5 Issue 8, August 2015.
25. Wang, G. G., Gandomi, A. H., Yang, X. S., and Alavi, A. H. (2016)A new hybrid method based on krill herd and cuckoo search for global optimization tasks, International Journal of Bio-Inspired Computation, Vol 8(5), pp. 286-299, 2016.
26. Pinelli, F ,Ginnotti, Pedreschi, D F., Nanni, M.,,: Proceedings of the 13th ACM SIGKDD, Trajectory Pattern Mining. In:, pp. 330–339 (2007).
27. Kuijpers and W.Othman : “Trajectory Databases- Data Models,Uncertainty and Complete Query Languages.”, ICDT, SpringerVerlag, LNCS vol. 4353, pp. 224-238, 2007.
28. Lee, J., Han, J., Li, X., Gonzalez, H.: ACM, VLDB, New Zealand 08, Tra Class- Trajectory classification Using Hierarchical Region Based and Trajectory based Clustering. In:, pp. 1081–1094 .
29. Jacinto, A. T.Mário and S. Jorge: IEEE Transactions on Image Processing Trajectory Classification Using Switched Dynamical Hidden Markov Models”, pp. 1338-1348, May 2010,Vol. 19, NO. 5.
30. Chen X, Pang J, Xue R. Transactions on the Web (TWEB), Constructing and comparing user mobility profiles, ACM. October 2014, , Article No. 21 ,Volume 8 Issue 4.
31. Tseng VS , Lee W-C ,Jia-Ching Ying J, Intelligent Mobile Knowledge Discovery and Management Systems, Special Issue , Special Section , Web Mining version , ACM, Transactions on Intelligent Systems and Technology (TIST) Mining geographic-temporal-semantic patterns in trajectories for location prediction - , Volume 5 Issue 1, December 2013, Article No. 2 .
32. A., Wang, Q , Yu ,.L., Rundensteiner,. At el. : In: ACM KDD, Detecting moving object outliers in massive-scale trajectory streams., pp. 422–431 (2014) .
33. Wang, G. G., Gandomi, Springer , Cham (2017): an intelligent QoS-aware autonomic Resource management approach for cloud computing, 49 -56..
34. Chen X, Kordy P , Lu R , Pang J. 2014. Machine Learning and Knowledge Discovery in Databases MinUS: Mining User Similarity with Trajectory Patterns., European Conference, ECML PKDD , Nancy, France,. Proceedings, Part III, Pages 436-439 , September 15-19, 2014
35. K. Sharma, O. P. Vyas, S. Scheider and A. Akasapu: “Nearest Neighbour Classification for Trajectory Data”, ITC 2010, Springer LNCS CCIS 101, pp. 180–185, 2010J. K.Wu, T. Tan , K F Loe at al. IEEE transaction .Image process. Spatiotemporal video Segmentation based on graphical model pp.937-947, July 2005.

36. Ahuja et al. IEEE Trans. Circuit System Video Technology, Inergration of frequency and space for multiple motion object segmentation, May 2008,vol.18 ,no. 5, pp 175-180.
37. A. Kuliko, Sun WW, RC Song et al., That Classification and Prediction Methods from Statistics, An analytical method for multi-class molecular cancer classification. SIAM Rev 45:726–734, 2014.
38. JN Yuan ,Y Zheng et al. IEEE T knowl Data Engineering ,Online discovery of gathering pattern ,2014. Pp 1974-1988.
39. Weibel , Weiskope, Speackmann, B. F., Safi K., , DE Wengh et al. N. V., an interdisciplinary review. *Movement Ecology*, R. Analysis and visualisation of movement: , 1–24. doi:10.1186/s40462-015-0032-y, 2016.
40. Guo, D ,Kupper, J. A., ZHU, X., AND. *Geographical Analysis* 48, 3 , Quantifying animal trajectories using spatial aggregation and sequence analysis – a case study of differentiating rajectories of multiple species., 275–291.doi:10.1111/gean.12098,2016.
41. Horton , M., Estepean, J , at el. . *19th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining* (2015), Automated classification of passing in football. In *PAKDD 2015*: Springer International Publishing, pp. 319–330. doi:10.1007/978-3-319-18032-8_25.

Figures

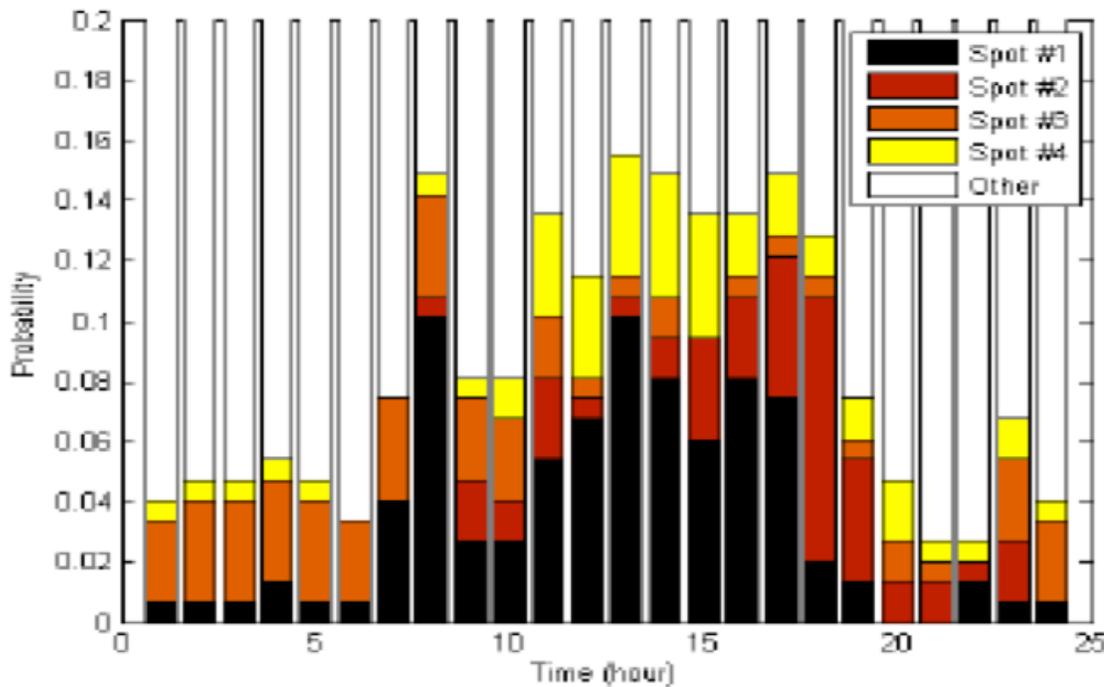


Figure 1

Periodic data movement pattern generation

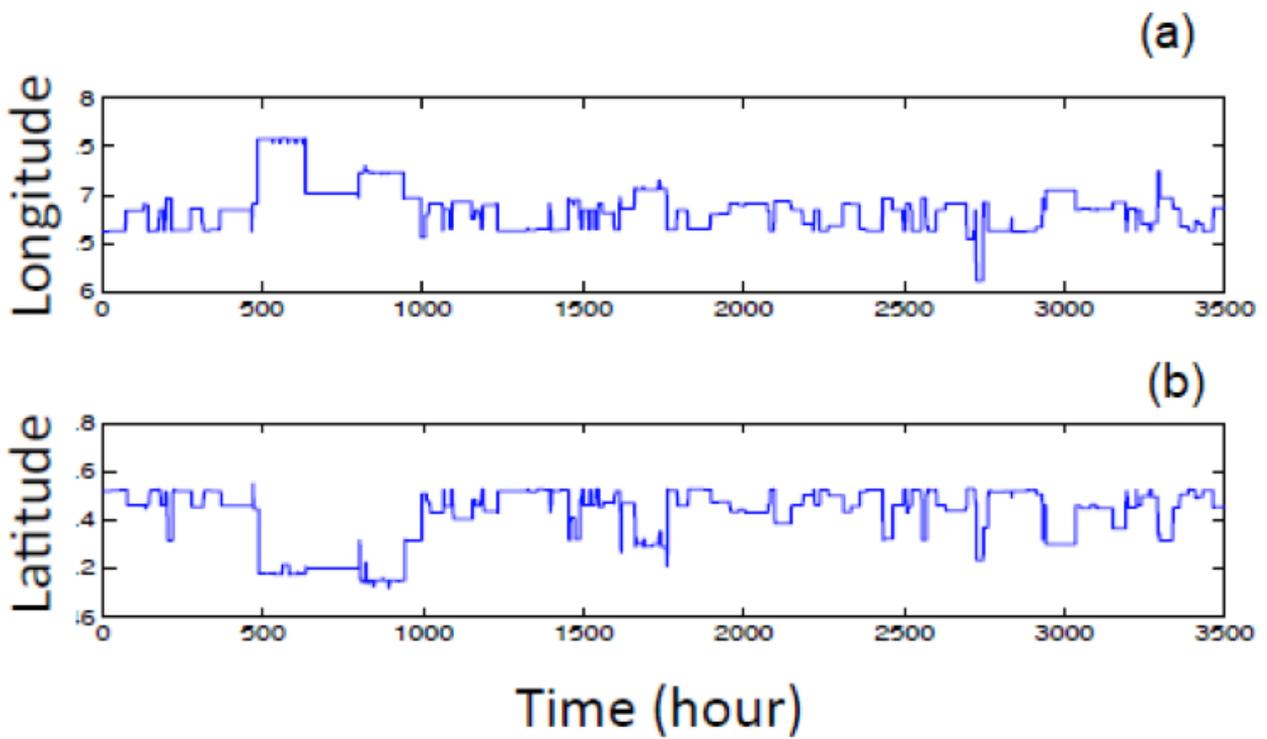


Figure 2

Trajectory data monthly generation prediction (movement with time): a. Dandkaran area Hurricane data measure. b. Dandkaran area Massless data measure.

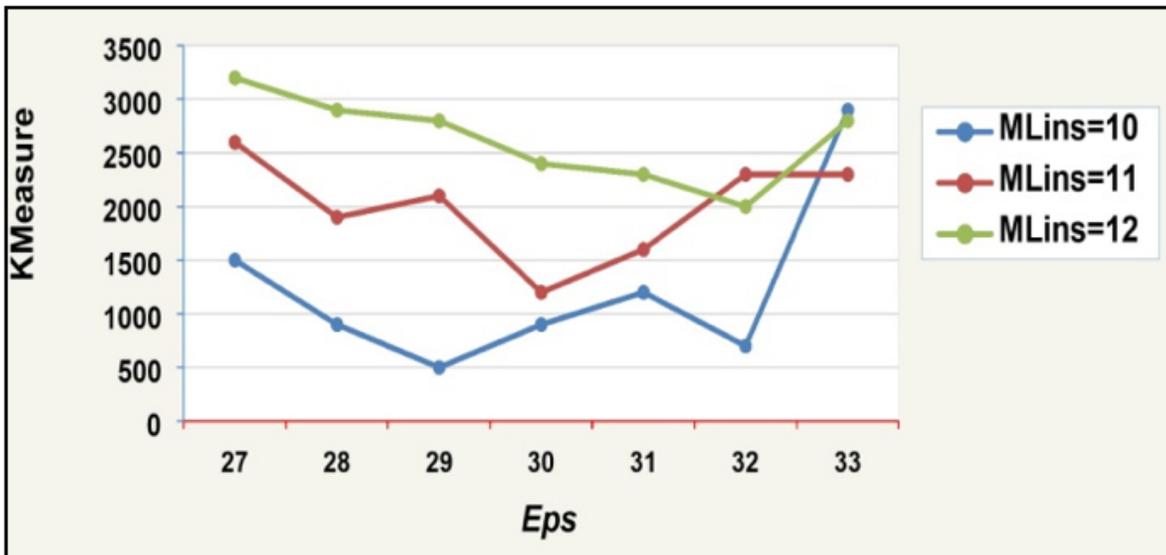


Figure 3

Hurricane data measure

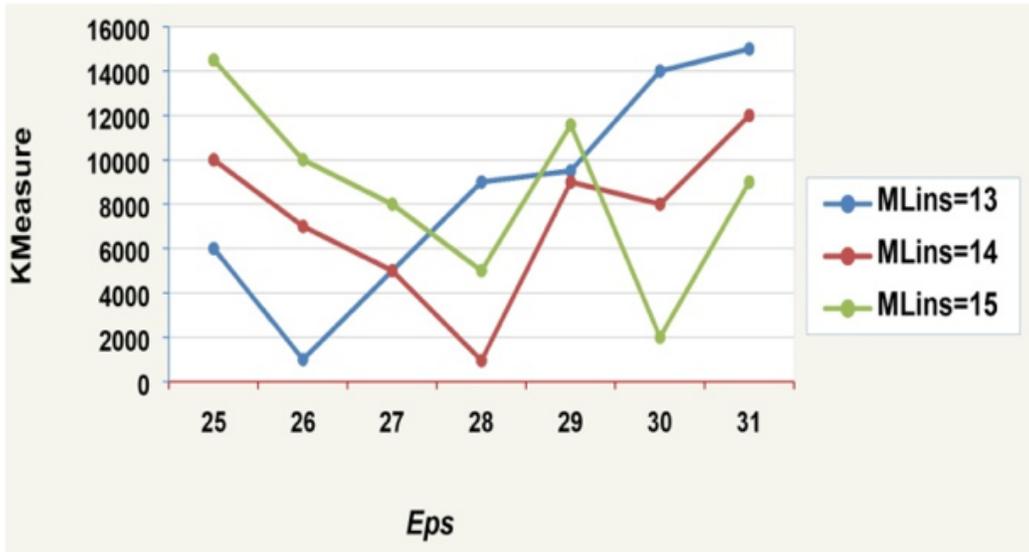


Figure 4

Massless data measure

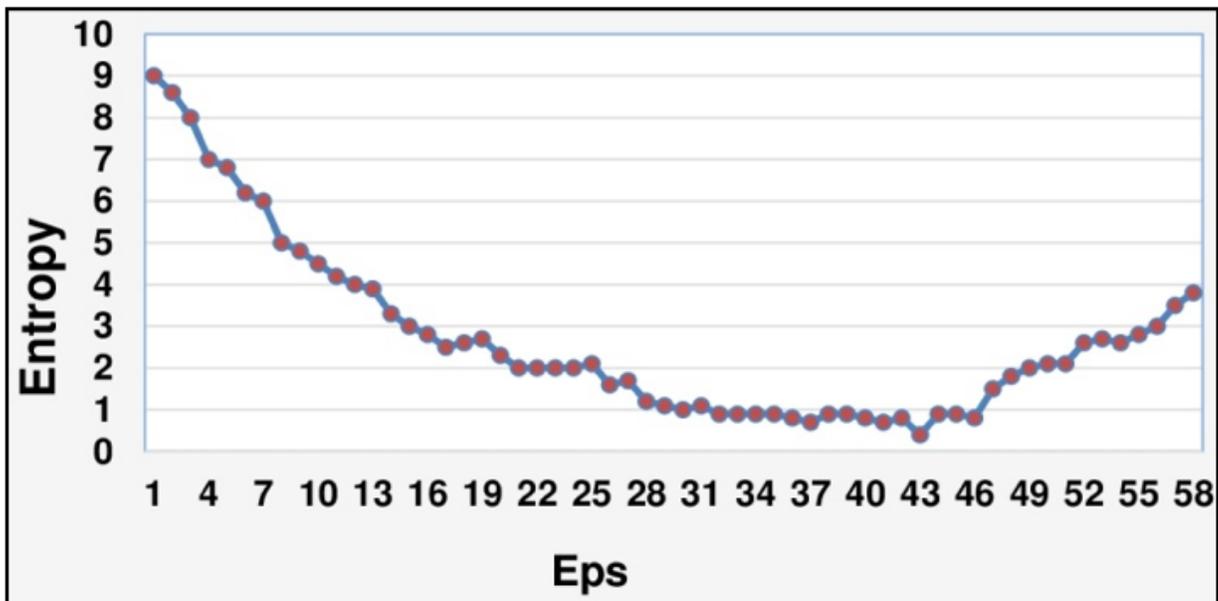


Figure 5

Hurricane data measure entropy

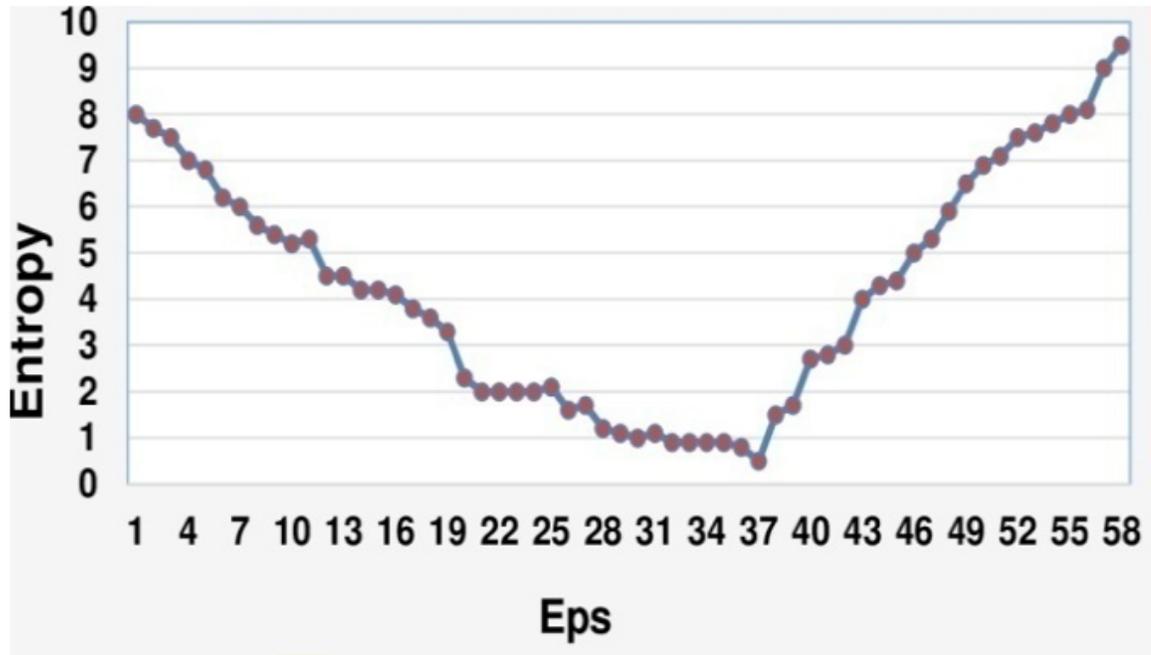


Figure 6

Massless data measure entropy