Supplementary Material

Section A. Web Links to Benchmark Datasets

Web links to the 14 benchmark datasets used for numerical study are given in Table S1.

Table S1. Web Links to Benchmark Datasets for Numerical Study

Dataset	Web link
R15	http://cs.joensuu.fi/sipu/datasets/
AG	http://cs.joensuu.fi/sipu/datasets/
S1	http://cs.joensuu.fi/sipu/datasets/
S2	http://cs.joensuu.fi/sipu/datasets/
S3	http://cs.joensuu.fi/sipu/datasets/
S4	http://cs.joensuu.fi/sipu/datasets/
AB	http://archive.ics.uci.edu/ml/datasets/Abalone
SP	https://archive.ics.uci.edu/ml/datasets/Spambase
CG	https://archive.ics.uci.edu/ml/datasets/Cardiotocography
SPF	https://archive.ics.uci.edu/ml/datasets/Steel+Plates+Faults
MF	https://archive.ics.uci.edu/ml/datasets/Multiple+Features
PD	https://archive.ics.uci.edu/ml/datasets/PenBased+Recognition+of+Handwritten+Digits
WQ	https://archive.ics.uci.edu/ml/datasets/wine+quality
OD	https://archive.ics.uci.edu/ml/datasets/Occupancy+Detection+
MNIST	http://yann.lecun.com/exdb/mnist/
FMNIST	https://www.kaggle.com/zalando-research/fashionmnist

Section B. Parameter Settings of Clustering Algorithms

Table S2. Parameter Settings of Clustering Algorithms Used in Numerical Examples

Algorithm	Free Parameter(s)	Experimental Setting
FCM	1) number of clusters, c; 2) fuzzy weighting exponent, m;	1) $c = T$; 2) $m = 2$;
MAL	1) number of clusters, c; 2) fuzzy weighting exponent, m; 3) model parameter, λ; 4) penalty parameter, ρ;	1) $c = T$; 2) $m = 2$; 3) $\lambda = 1$; 4) $\rho = 1$ [1];
KM	number of clusters, k;	$k = T$;
DBS	1) cluster radius, r; 2) minimum number of data samples within the radius, k;	1) the value of the knee point of the sorted *k-dist* graph; 2) $k = 4$;
MS	1) bandwidth, p; 2) kernel function type;	1) $p = 0.3$; 2) Gaussian kernel;
SUB	initial cluster radius, r;	$r = 0.3$ [2]
NMI	1) bandwidth sequence; 2) prior data distribution model;	1) $[0.1\hat{\sigma}, 0.2\hat{\sigma}, \dots, 2.0\hat{\sigma}]$; $\hat{\sigma}$ is the maximum value of the standard deviations calculated from each attribute/feature of the data [3]; 2) Gaussian distribution;
AP	1) maximum number of iterative refinements; 2) cumulative number of iterations for monitoring the exemplar decisions; 3) dampening factor, λ;	1) 200; 2) 20; 3) $\lambda = 0.5$ [4];
GDD	none	-
CLA	number of neighbours, k	$k = \begin{cases} \lceil 0.1N \rceil & if\ N \leq 100 \\ 10 + \lceil 0.02N \rceil & else \end{cases}$ [5]
LGC	1) number of nearest neighbours, k; 2) initial momentum, IM; 3) critical centrality value, $\widehat{CE}$;	1) $k = \max([5, \min([15, \lceil 0.015N \rceil])])$; 2) $IM = 10$; 3) $\widehat{CE} = 0.6$ [5];

ADP	none	-
OSKM	1) number of clusters, k; 2) initial learning rate, η_0; 3) final learning rate, η_f;	1) $k = T$; 2) $\eta_0 = 1.0$; 3) $\eta_f = 0.01$ [6];
ECL	1) initial cluster radius, r; 2) threshold for adding new cluster, ε; 3) centre updating ratio, ρ;	1) $r = 0.5$; 2) $\varepsilon = e^{-1}$; 3) $\rho = 0.5$ [7];
ELM	kernel bandwidth, r;	$r = 0.15$ [8];
OCA	1) probability value, P; 2) minimum weight, p;	1) $P = 0.99$; 2) $p = 0.1$ [9].

Section C. Clustering Quality Indices

1) Adjusted Rand index (ARI) [10]

Suppose that $\boldsymbol{y} = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_n\}$ and $\boldsymbol{z} = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_m\}$ are two different partitions of $\{\boldsymbol{x}\}_K$, the overlaps between and can be summarized by the contingency table given by Fig. S1, where $s_{i,j}$ denotes the number of comment elements of $\boldsymbol{y}_i$ and $\boldsymbol{z}_j$: $s_{i,j} = |\boldsymbol{y}_i \cap \boldsymbol{z}_j|$.

$\boldsymbol{y}$ \ $\boldsymbol{z}$	$\boldsymbol{z}_1$	$\boldsymbol{z}_2$	…	$\boldsymbol{z}_m$	sum
$\boldsymbol{y}_1$	$s_{1,1}$	$s_{1,2}$	…	$s_{1,m}$	$s_{1,*}$
$\boldsymbol{y}_2$	$s_{2,1}$	$s_{2,2}$	…	$s_{2,m}$	$s_{2,*}$
…	…	…	$\ddots$	…	…
$\boldsymbol{y}_n$	$s_{n,1}$	$s_{n,2}$	…	$s_{n,m}$	$s_{n,*}$
sum	$s_{*,1}$	$s_{*,2}$	…	$s_{*,m}$	

Fig. S1. Contingency table of the two partitions

ARI is defined by Eqn. (S1) as follows.

$$ARI = \frac{\sum_{j=1}^{n}\sum_{i=1}^{m}\binom{s_{i,j}}{2} - \left[\sum_{i=1}^{m}\binom{s_{*,i}}{2}\sum_{j=1}^{n}\binom{s_{j,*}}{2}\right]/\binom{K}{2}}{\frac{1}{2}\left[\sum_{i=1}^{m}\binom{s_{*,i}}{2}+\sum_{j=1}^{n}\binom{s_{j,*}}{2}\right] - \left[\sum_{i=1}^{m}\binom{s_{*,i}}{2}\sum_{j=1}^{n}\binom{s_{j,*}}{2}\right]/\binom{K}{2}} \tag{S1}$$

2) Calinski Harabasz index (CHI) [11]

CHI is defined by Eqn. (S2):

$$CHI = \frac{(K-C)\sum_{i=1}^{C} S_i \|\boldsymbol{m}_i - \bar{\boldsymbol{x}}\|^2}{(C-1)\sum_{i=1}^{C}\sum_{j=1}^{S_i}\|\boldsymbol{m}_i - \boldsymbol{x}_{i,j}\|^2} \tag{S2}$$

where $\bar{\boldsymbol{x}}$ is the mean of data, namely, $\bar{\boldsymbol{x}} = \frac{1}{K}\sum_{i=1}^{K}\boldsymbol{x}_i$; $\boldsymbol{m}_i$ is the mean of all samples within the i^{th} cluster; $\boldsymbol{x}_{i,j}$ is the j^{th} sample of the i^{th} cluster; S_i is the number of data samples of the i^{th} cluster.

3) Davies-Bouldin index (DBI) [12]

DBI is defined as follows.

$$DBI = \frac{1}{C}\sum_{i=1}^{C}\max_{\substack{j=1,2,\ldots,C;\\ j\neq i}}\left(\frac{\frac{1}{S_i}\sum_{k=1}^{S_i}\|\boldsymbol{m}_i - \boldsymbol{x}_{i,k}\| + \frac{1}{S_j}\sum_{k=1}^{S_j}\|\boldsymbol{m}_j - \boldsymbol{x}_{j,k}\|}{\|\boldsymbol{m}_i - \boldsymbol{m}_j\|}\right) \tag{S3}$$

4) Silhouette coefficient (SC) [13]

To calculate SC, the silhouette value (SI) of every data sample, $\boldsymbol{x}_{i,j}$ is calculated as follows:

$$SI(\boldsymbol{x}_{i,j}) = \begin{cases} \frac{\beta(\boldsymbol{x}_{i,j}) - \alpha(\boldsymbol{x}_{i,j})}{\max([\alpha(\boldsymbol{x}_{i,j}), \beta(\boldsymbol{x}_{i,j})])} & if\ S_i > 1 \\ 0 & if\ S_i = 1 \end{cases} \tag{S4}$$

where $\alpha(\boldsymbol{x}_{i,j})$ and $\beta(\boldsymbol{x}_{i,j})$ are defined as:

$$\begin{cases} \alpha(\boldsymbol{x}_{i,j}) = \frac{1}{S_i - 1}\sum_{k=1}^{S_i}\left\|\boldsymbol{x}_{i,j} - \boldsymbol{x}_{i,k}\right\| \\ \beta(\boldsymbol{x}_{i,j}) = \min_{\substack{l=1,2,\dots,C;\\ l\neq i}} \frac{1}{S_i - 1}\sum_{k=1}^{S_l}\left\|\boldsymbol{x}_{i,j} - \boldsymbol{x}_{l,k}\right\| \end{cases} \tag{S5}$$

The average SI value per cluster is then obtained as ($i = 1,2,\dots,C$):

$$\overline{SI}_i = \frac{1}{S_i}\sum_{k=1}^{S_i} SI(\boldsymbol{x}_{i,k}) \tag{S6}$$

Finally, SC is obtained as the maximum average SI value of the clusters in the data space:

$$SC = \max_{i=1,2,\dots,C}(\overline{SI}_i) \tag{S7}$$

Section D. Sensitivity Analysis of Proposed Algorithm

Table S3. Static Data Clustering Performance Demonstration with Different Levels of Granularity

G	Dataset	Measure					
		C	ARI	CHI	DBI	SC	t_{exe}
2	R15	8	0.264	759.358	0.351	0.780	0.018
3		8	0.264	760.074	0.349	0.781	0.012
4		15	0.986	4846.116	0.315	0.900	0.013
5		16	0.974	4682.205	0.414	0.869	0.014
6		29	0.818	3345.041	0.827	0.605	0.014
7		100	0.330	3237.396	0.710	0.515	0.038
2	AG	4	0.744	1166.866	0.623	0.698	0.026
3		5	0.816	1085.225	0.561	0.696	0.021
4		8	0.736	1365.448	0.701	0.649	0.022
5		15	0.412	1658.009	0.774	0.610	0.022
6		27	0.255	1576.486	0.841	0.541	0.023
7		100	0.330	3237.396	0.710	0.515	0.038
2	S1	1	0.000	NaN	NaN	NaN	0.988
3		7	0.515	6555.791	0.680	0.668	1.013
4		15	0.986	22675.170	0.366	0.880	0.958
5		20	0.901	18323.770	0.829	0.731	1.023
6		91	0.441	11653.290	0.951	0.458	2.122
7		260	0.225	11401.350	0.779	0.464	6.422
2	S2	3	0.201	4406.222	0.838	0.560	0.897
3		8	0.504	5251.601	0.753	0.585	0.910
4		15	0.936	13505.920	0.465	0.801	0.899
5		18	0.919	11633.440	0.604	0.759	0.882
6		81	0.513	8059.474	0.855	0.554	1.305
7		267	0.209	7229.232	0.809	0.433	9.689
2	S3	2	0.110	4282.463	0.989	0.580	0.934
3		5	0.347	5126.302	0.876	0.581	0.926
4		12	0.609	6775.569	0.667	0.630	0.912
5		23	0.641	6344.410	0.862	0.600	0.951
6		82	0.364	4845.483	0.902	0.470	1.724
7		224	0.223	4208.581	0.747	0.483	5.113
2	S4	2	0.108	3244.414	1.131	0.526	0.941
3		5	0.313	4501.981	0.829	0.544	0.957
4		16	0.588	5463.927	0.731	0.632	0.933
5		36	0.446	4648.732	0.934	0.485	1.125
6		104	0.303	3643.270	0.822	0.469	2.576
7		276	0.177	3962.409	0.678	0.481	8.078

Table S4. Streaming Data Clustering Performance Demonstration with Different Numbers of Chunks

G	L	Dataset	Measure					
			C	ARI	CHI	DBI	SC	t_{exe}
4	1	AB	36	0.044	8822.611	0.485	0.709	0.727
	2		41	0.036	8285.478	0.482	0.722	0.390
	3		46	0.035	7299.781	0.493	0.721	0.261
	4		47	0.035	8255.470	0.474	0.715	0.218
	5		50	0.035	8947.589	0.475	0.713	0.158
5	1		83	0.027	9841.226	0.518	0.671	7.413
	2		92	0.033	12517.460	0.506	0.664	1.936
	3		98	0.034	11740.080	0.497	0.654	1.136
	4		104	0.035	12150.620	0.489	0.667	0.808
	5		107	0.033	12856.950	0.509	0.650	0.529
4	1	SB	184	0.100	22737.400	0.322	0.615	2.795
	2		132	0.107	1502.662	2.143	0.485	1.397
	3		152	0.114	6018.315	1.315	0.606	0.851
	4		155	0.136	4062.510	1.332	0.613	0.569
	5		162	0.131	5738.480	1.841	0.614	0.494
5	1		287	0.044	30434.440	0.322	0.489	3.950
	2		220	0.149	4544.743	1.364	0.491	1.413
	3		242	0.075	11768.820	0.757	0.523	0.750
	4		238	0.147	4490.374	0.686	0.506	0.750
	5		252	0.078	5034.233	0.985	0.539	0.542
4	1	PD	45	0.361	1297.889	1.720	0.252	4.504
	2		28	0.516	1391.596	1.585	0.392	2.304
	3		39	0.504	1203.205	1.560	0.367	1.605
	4		39	0.540	1193.203	1.536	0.372	1.231
	5		48	0.536	1044.625	1.529	0.354	0.986
5	1		154	0.168	666.815	1.766	0.164	5.124
	2		116	0.516	521.650	1.391	0.311	2.860
	3		179	0.352	452.573	1.402	0.255	1.800
	4		188	0.425	392.472	1.358	0.265	1.403
	5		226	0.351	369.892	1.325	0.240	1.196
4	1	WQ	22	0.001	6206.402	0.683	0.489	2.541
	2		24	0.000	4661.316	0.703	0.474	1.028
	3		28	0.000	4487.105	0.720	0.466	0.669
	4		30	0.001	4671.105	0.714	0.440	0.489
	5		32	-0.001	4208.644	0.693	0.445	0.362
5	1		42	0.002	4518.470	0.681	0.468	2.113
	2		42	0.000	2248.599	0.711	0.335	1.052
	3		52	0.000	3270.815	0.685	0.405	0.671
	4		60	0.001	3355.187	0.734	0.393	0.485
	5		62	0.001	3718.894	0.756	0.386	0.386
4	1	OD	89	0.126	59096.400	0.536	0.709	35.849
	2		106	0.132	49365.410	0.578	0.678	18.131
	3		111	0.134	47806.160	0.603	0.673	10.542
	4		126	0.119	50974.690	0.608	0.664	9.190
	5		131	0.121	50120.910	0.622	0.666	6.778
5	1		186	0.109	56560.040	0.616	0.673	44.025
	2		135	0.221	18030.080	0.745	0.469	22.746
	3		218	0.098	49300.500	0.709	0.612	14.241
	4		184	0.157	24054.140	0.753	0.409	11.016
	5		209	0.152	27088.020	0.730	0.543	7.785

Section E. Performance Ranking of Clustering Algorithms

Performance ranks per criterion of all the clustering algorithms on the benchmark datasets during the numerical experiments are reported in Tables S5-S7. Note that average ranks are reported if the obtained values of a certain criterion by two or more algorithms on a particular dataset are the same.

Table S5. Static Data Clustering Performance Ranks on Benchmark Synthetic Datasets per Criterion

Algorithm	Dataset	Measure				Dataset	Measure			
		ARI	CHI	DBI	SC		ARI	CHI	DBI	SC
AFC ($G = 4$)	R15	5.5	4	3.5	3.5	AG	9	4	8	2
AFC ($G = 5$)		8	7	9	8		12	1	11	8
FCM		7	8	8	7		10	7	14	6
MAL		10	10	12	11		4	10	5	4
KM		9	9	10	9		8	5	9	1
DBS		3	2	1	1.5		3	13	2	12
MS		13.5	13.5	13.5	13.5		11	14	10	14
SUB		12	12	7	10		6	6	7	5
NMI		13.5	13.5	13.5	13.5		7	8	6	7
AP		5.5	6	6	5.5		14	3	13	11
GDD		11	11	11	12		5	12	4	13
CLA		3	5	3.5	5.5		1	9	1	3
LGC		3	3	3.5	3.5		2	11	3	9
ADP		1	1	3.5	1.5		13	2	12	10
AFC ($G = 4$)	S1	3	2	1	1.5	S2	1	1	1	1
AFC ($G = 5$)		8	6	12	7		3.5	6	5	6
FCM		5	5	4	4		3.5	3	3	3
MAL		9	9	10	9		9	9	12	10
KM		7	7	6	6		6	4	4	4
DBS		6	8	11	10		8	11	13	11
MS		13	12	13	12		12	12	11	12
SUB		11	10	8	8		10	8	7	8
NMI		12	11	9	11		11	10	9	9
AP		14	14	14	14		14	14	14	14
GDD		10	13	7	13		13	13	10	13
CLA		1.5	3	2.5	3		2	2	2	2
LGC		4	4	5	5		7	7	8	7
ADP		1.5	1	2.5	1.5		5	5	6	5
AFC ($G = 4$)	S3	8	4	4	3	S4	4	4	5	2
AFC ($G = 5$)		6	7	8	6		8	6	12	7
FCM		2	1	2.5	1		1.5	1	2	1
MAL		9	10	12	10		10	10	13	11
KM		4	5	6	5		3	3	4	3
DBS		12	11	10	11		13	12	11	12
MS		13.5	13.5	13.5	13.5		12	11	8	10
SUB		10	8	5	8		9	7	6	6
NMI		1	2	2.5	2		1.5	2	3	4
AP		13.5	13.5	13.5	13.5		14	14	14	14
GDD		11	12	11	12		11	13	1	13
CLA		3	3	1	4		5	8	7	9
LGC		7	9	9	9		7	9	10	8
ADP		5	6	7	7		6	5	9	5

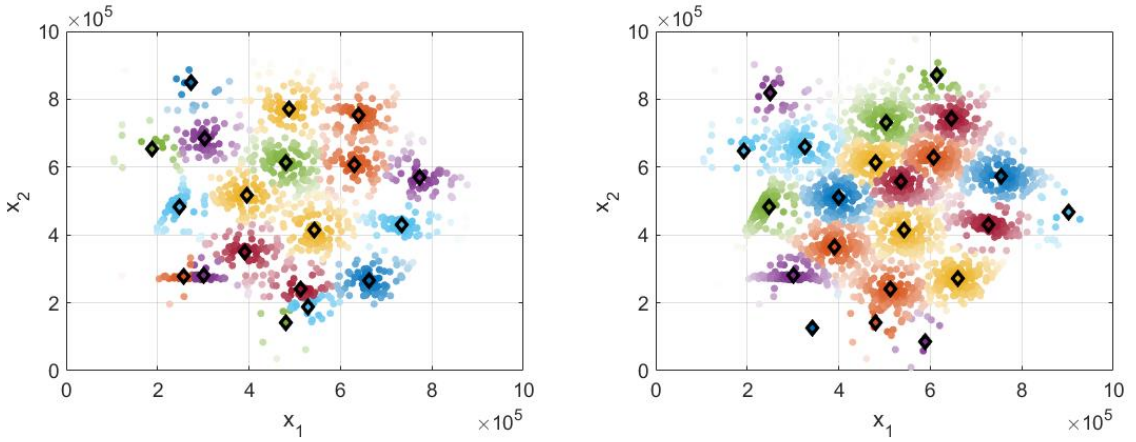
Table S6. Static Data Clustering Performance Ranks on Benchmark Real-World Datasets per Criterion

Algorithm	Dataset	Measure				Dataset	Measure			
		ARI	CHI	DBI	SC		ARI	CHI	DBI	SC
AFC ($G = 4$)	AB	6	3	2	3	SB	3	2	2.5	5
AFC ($G = 5$)		11	2	3	6		8	1	2.5	6
FCM		2	5	6	5		7	3	5	2
MAL		1	10	12	11		10	10	10	7
KM		3.5	7	7	7		11	4	4	1
DBS		10	4	10	2		12	13.5	13.5	13.5
MS		12.5	13	11	13		13	12	11	12
SUB		3.5	11	13	12		1	11	12	11
NMI		12.5	12	1	1		9	8	1	4
AP		14	14	14	14		14	13.5	13.5	13.5
GDD		7	6	4	9		6	7	7	8
CLA		8.5	1	8	4		4	6	9	10
LGC		8.5	9	9	10		2	9	8	9
ADP		5	8	5	8		5	5	6	3
AFC ($G = 4$)	CG	3	6	4	7	SPF	5	2	2	2
AFC ($G = 5$)		8	8	3	6		7.5	1	3	4
FCM		2	2	9	8		10	5	5	6
MAL		4	4	12	10		1	12	12	10
KM		1	1	6	4		9	4	4	5
DBS		9	12	7	12		2	9	10	12
MS		14	13.5	13.5	13.5		13.5	13.5	13.5	13.5
SUB		10	10	11	11		11	10	11	9
NMI		5	11	1	9		12	8	1	1
AP		7	3	5	5		13.5	13.5	13.5	13.5
GDD		12	7	10	1		4	11	8	11
CLA		13	13.5	13.5	13.5		3	6	7	7
LGC		11	9	8	2		7.5	7	9	8
ADP		6	5	2	3		6	3	6	3
AFC ($G = 4$)	MF	3	3	6	5	PD	7	5	8	6
AFC ($G = 5$)		8	7	7	7		10	8	9	8
FCM		5	1	2	1		5	3	13	3
MAL		9	10	10	9		3	2	12	5
KM		4	2	3	2		2	1	4	1
DBS		10	9	9	10		4	10	7	10
MS		12.5	12.5	12.5	12.5		6	12	1	12
SUB		12.5	12.5	12.5	12.5		11	11	10	11
NMI		12.5	12.5	12.5	12.5		14	14	14	14
AP		2	5	5	6		12	9	6	7
GDD		12.5	12.5	12.5	12.5		13	13	2	13
CLA		7	8	8	8		9	6	11	9
LGC		1	4	1	3		1	4	5	4
ADP		6	6	4	4		8	7	3	2
AFC ($G = 4$)	WQ	8	4	4	3	OD	9	1	2	5
AFC ($G = 5$)		6	5	3	5		11	2	4	6
FCM		6	2	6	2		2	5	6	2
MAL		1.5	9	13	9		12	11	12	7
KM		4	1	5	1		3	4	5	1
DBS		13	10	8	12		6	10	9	12
MS		10	13	9	13		13	12	11	11
SUB		1.5	6	12	8		10	8	13	9
NMI		10	8	1	7		1	7	1	4
AP		14	14	14	14		14	14	14	14
GDD		6	11	11	11		7	13	10	13
CLA		3	7	2	6		8	9	8	10
LGC		12	12	10	10		5	6	7	8
ADP		10	3	7	4		4	3	3	3

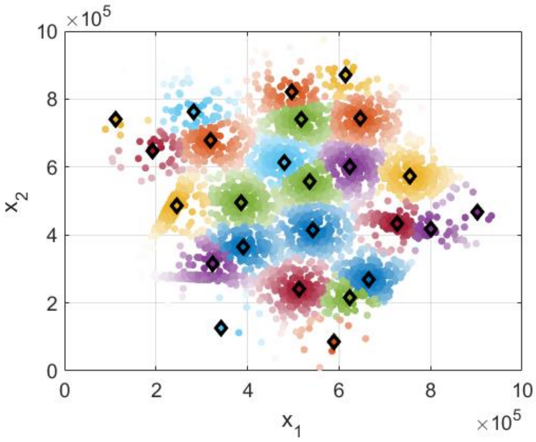
Table S7. Streaming Data Clustering Performance Ranks on Benchmark Real-World Datasets per Criterion

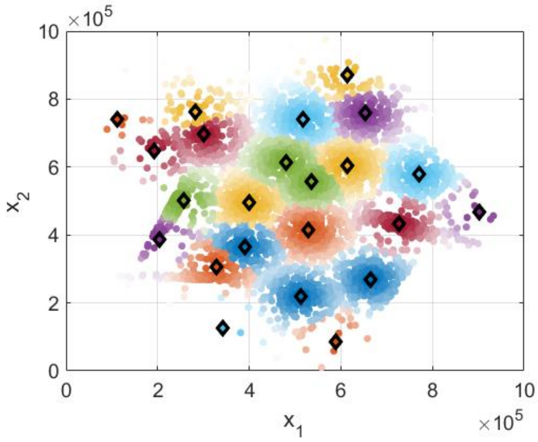
Algorithm	Dataset	Measure				Dataset	Measure			
		ARI	CHI	DBI	SC		ARI	CHI	DBI	SC
AFC ($G = 4$)	AB	5	2	1	1	SB	2	1	5	1
AFC ($G = 5$)		6	1	2	2		4	2	2	3
OKM		2	5	6	5		6	7	3	6
EC		3	4	3	4		1	4	4	4
ELM		7	7	5	7		7	5	6	7
OCA		1	6	7	6		5	6	7	5
ADP		4	3	4	3		3	3	1	2
AFC ($G = 4$)	CG	3	3	2	3	SPF	3	2	1	1
AFC ($G = 5$)		5	5	1	2		5	1	2	2
OKM		1.5	4	6	6		1	6	6	7
EC		1.5	1	4	4		4	4	4	4
ELM		6	7	5	7		6	5	5	6
OC		7	6	7	5		7	7	7	5
ADP		4	2	3	1		2	3	3	3
AFC ($G = 4$)	MF	1	2	3	2	PD	1	3	3	1
AFC ($G = 5$)		2	4	1	3		2	5	1	3
OKM		5	5	5	5		4	2	6	5
EC		4	3	4	4		3	1	5	4
ELM		7	6.5	6.5	6.5		7	6	4	7
OCA		6	6.5	6.5	6.5		6	7	7	6
ADP		3	1	2	1		5	4	2	2
AFC ($G = 4$)	WQ	7	2	1	1	OD	6	1	1	2
AFC ($G = 5$)		6	4	2	2		5	3	3	3
OKM		1	6	6	6		4	5	7	5
EC		2	3	4	5		1	4	4	4
ELM		4.5	7	5	7		2	7	6	7
OCA		3	5	7	4		7	6	5	6
ADP		4.5	1	3	3		3	2	2	1

Section F. Demonstration of Streaming Data Clustering Process of Proposed Algorithm



(a) Clustering result obtained with chunk 1 (b) Clustering result obtained with chunks 1 & 2





(c) Clustering result obtained with chunks 1 & 2 & 3 (d) Clustering result obtained with all four chunks

Fig. S2. Evolution of the clustering outcome of AFC with $G = 4$ and $L = 4$ (dots – data samples; diamonds – cluster medoids)

Table S8. Evolution of Clustering Quality Criteria Values over Online Streaming Data Clustering Process

Processed Chunks	Measure				
	C	ARI	CHI	DBI	SC
1	19	0.576	1335.314	0.757	0.609
1,2	22	0.577	2475.506	0.769	0.611
1,2,3	25	0.560	3479.851	0.798	0.574
1,2,3,4	23	0.582	4756.509	0.761	0.608

References

[1] L. Guo, L. Chen, X. Lu, and C. L. P. Chen, “Membership affinity lasso for fuzzy clustering,” *IEEE Trans. Fuzzy Syst.*, vol. 28, no. 2, pp. 294–307, 2020.

[2] S. L. Chiu, “Fuzzy model identification based on cluster estimation,” *J. Intell. Fuzzy Syst.*, vol. 2, no. 3, pp. 267–278, 1994.

[3] J. Li, S. Ray, and B. G. Lindsay, “A nonparametric statistical approach to clustering via mode identification,” *J. Mach. Learn. Res.*, vol. 8, no. 8, pp. 1687–1723, 2007.

[4] B. J. Frey and D. Dueck, “Clustering by passing messages between data points,” *Science (80-.).*, vol. 315, no. 5814, pp. 972–976, 2007.

[5] Z. Wang *et al.*, “Clustering by local gravitation,” *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1383–1396, 2018.

[6] S. Zhong, “Efficient online spherical k-means clustering,” in *Proceedings of the International Joint Conference on Neural Networks*, 2005, pp. 3180–3185.

[7] P. Angelov, “An approach for fuzzy rule-base adaptation using on-line clustering,” *Int. J. Approx. Reason.*, vol. 35, no. 3, pp. 275–289, 2004.

[8] R. Dutta Baruah and P. Angelov, “Evolving local means method for clustering of streaming data,” in *IEEE International Conference on Fuzzy Systems*, 2012, pp. 10–15.

[9] M. Chenaghlou, M. Moshtaghi, C. Leckie, and M. Salehi, “Online clustering for evolving data streams with online anomaly detection,” in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2018, pp. 508–521.

[10] L. Hubert and P. Arabie, “Comparing partitions,” *J. Classif.*, vol. 2, no. 1, pp. 193–218, 1985.

[11] T. Caliński and J. Harabasz, “A dendrite method for cluster analysis,” *Commun. Stat. Methods*, vol. 3, no. 1, pp. 1–27, 1974.

[12] D. L. Davies and D. W. Bouldin, “A cluster separation measure,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 2, pp. 224–227, 1979.

[13] P. J. Rousseeuw, “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis,” *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, 1987.