The authors would like to make the following corrections to this paper [1]:

(1) We corrected the subscript error in Definition 2.

Replacing:

**Definition 2.** The minimum distance between a cell $g_i$ and a cluster center $c_j$ for numeric attributes, denoted $d_{\text{min}}(g_i, c_j)$, is:

$$d_{\text{min}}(g_i, c_j) = \sqrt{\sum_{r_i=1}^{m_i} |o_i - r_i|^2},$$

(5)

where $r_i = \begin{cases} s_i & \text{if } o_i < s_i \\ t_i & \text{if } o_i > t_i \\ a_i & \text{otherwise} \end{cases}$. 

with:

**Definition 2.** The minimum distance between a cell $g_i$ and a cluster center $c_j$ for numeric attributes, denoted $d_{\text{min}}(g_i, c_j)$, is:

$$d_{\text{min}}(g_i, c_j) = \sqrt{\sum_{r_i=1}^{m_i} |c_{ji} - r_i|^2},$$

(5)

where $r_i = \begin{cases} s_i & \text{if } c_{ji} < s_i \\ t_i & \text{if } c_{ji} > t_i \\ a_i & \text{otherwise} \end{cases}$. 

(2) We corrected the subscript error in Definition 3.

Replacing:

**Definition 3.** The maximum distance between a cell $g_i$ and a cluster center $c_j$ for numeric attributes, denoted $d_{\text{max}}(g_i, c_j)$, is:

$$d_{\text{max}}(g_i, c_j) = \sqrt{\sum_{r_i=1}^{m_i} |p_i - r_i|^2},$$

(6)

where $r_i = \begin{cases} t_i, & p_i \leq \frac{s_i + t_i}{2} \\ s_i, & \text{otherwise} \end{cases}$.
with:

\[ d_{\text{max}}(g_i, c_j) = \sqrt{\sum_{i=1}^{m_i} |c_{ji} - r_i|^2}, \]  

(6)

where \( r_i = \begin{cases} t_i, & c_{ji} \leq s_i + t_i \\ s_i, & \text{otherwise} \end{cases} \)

(3) We corrected the order of some commands in Algorithm 1.

Replacing:

1: \( C[\text{ }] \leftarrow \emptyset \) // k cluster centers
2: Randomly choosing \( k \) object, and assigning it to \( C \).
3: while IsConverged() do
4: \( d_{\text{min}}, d_{\text{max}} \leftarrow \text{Calc}(g, C) \)
5: \( d_{\text{min max}} \leftarrow \text{min}(d_{\text{max}}[\text{ }]) \)
6: for each cell \( g \) in \( G \) with:

   1: \( C[\text{ }] \leftarrow \emptyset \) // k cluster centers
   2: Randomly choosing \( k \) object, and assigning it to \( C \).
   3: while IsConverged() do
   4: \( d_{\text{min}}, d_{\text{max}} \leftarrow \text{Calc}(g, C) \)
   5: \( d_{\text{min max}} \leftarrow \text{min}(d_{\text{max}}[\text{ }]) \)

(4) We corrected the order of some commands in Algorithm 2.

Replacing:

1: \( C[k] \leftarrow \emptyset \) // \( k \) cluster center
2: Randomly choosing \( k \) object, and assigning it to \( C \).
3: while IsConverged() do
4: \( d_{\text{min}}, d_{\text{max}} \leftarrow \text{Calc}(g, C) \)
5: \( d_{\text{min max}} \leftarrow \text{min}(d_{\text{max}}[\text{ }]) \)
6: for each cell \( g \) in \( G \) with:

   1: \( C[\text{ }] \leftarrow \emptyset \) // \( k \) cluster centers
   2: Randomly choosing \( k \) object, and assigning it to \( C \).
   3: while IsConverged() do
   4: \( d_{\text{min}}, d_{\text{max}} \leftarrow \text{Calc}(g, C) \)
   5: \( d_{\text{min max}} \leftarrow \text{min}(d_{\text{max}}[\text{ }]) \)

(5) We corrected the analysis of complexity.

Replacing:

However, our proposed algorithms based on heuristic techniques (KCP and KBP) can reduce
the number of objects to be computed, \( n' \leq n \), and the number of dimensions to be computed, \( d' \leq d \), respectively. Therefore, the time complexities of our proposed algorithms are \( O(n'kd') \), \( n' \leq n \) and \( d' \leq d \).
For space complexity, KCP requires $O(nd)$ to store the entire dataset, $O(gd)$ to store the start point vector $S$ and the end point vector of each cell, $O(km_c)$ to store the frequency of categorical data in each cluster and $O(kd)$ to store cluster centers, where $m_c$ the number of categorical attributes. Additionally, KBP requires $O(gm_r)$ to store the frequency of categorical data in each cell, where $g$ is the number of cells. Therefore, the space complexities of KCP and KBP are $O(nd + gd + km_c + kd + gm_r)$, respectively.

However, our proposed algorithms based on heuristic techniques (KCP and KBP) can reduce the number of cluster centers to be computed, $k' \leq k$, and the number of dimensions to be computed, $d' \leq d$. Therefore, the time complexities of our proposed algorithms are $O(nkd''i), k' \leq k$, and $d' \leq d$.

For space complexity, KCP requires $O(nd)$ to store the entire dataset, $O(gm_r)$, where $g$ is the number of cells, to store the start point vector $S$ and the end point vector $T$ of each cell, $O(km_c)$, where $t$ is the number of categorical data, to store the frequency of categorical data in each cluster and $O(kd)$ to store cluster centers. KBP requires $O(gtm_c)$ to store the bitmap index in each cell, where $g$ is the number of cells. Therefore, the space complexities of KCP and KBP are $O(nd + gm_r + ktm_c + kd)$ and $O(nd + gm_r + ktm_c + kd + gtm_c)$, respectively.

(6) We corrected the description of Figure 12.

Replacing:

**Figure 12.** Effect of the number of clusters (numeric data and categorical data are on uniform distribution).

with:

**Figure 12.** Effect of the number of clusters (numeric data are uniformly distributed and the distribution of categorical data is skewed).

The authors would like to apologize for any inconvenience caused to the readers by these changes. The manuscript will be updated and the original will remain available from the article webpage.

Reference