Research on Railway Emergency Rescue Decision-Making Method Based on Clustering and SVD Algorithm

Jing Zuo¹,², Bin Shuai¹,*, Min LV¹ and Wen-cheng Huang¹

¹College of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China
²School of Automation & Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

*Corresponding author

Keywords: Railway transportation, SVD, Clustering, Emergency decision-making.

Abstract. This paper applied a new method based on clustering and SVD to railway emergency rescue decision-making, for it can provide an applicable, intelligent and efficient decision-making method for railway emergency rescue with high-dimensional data. The method SVD has been used to process the feature attributes of cases according to feature attribute extraction of emergency cases. Then the paper structured a railway emergency rescue decision-making model based on clustering method. Following it, a specific example was used to illustrate the railway unconventional emergency decision-making process based on clustering and SVD method, and the results showed the clustering and SVD method can meet the railway emergency decision-making requirement to a large extent. This paper made up the shortage of speculative knowledge in this aspect, offered a new method and idea in the railway emergency decision support system intelligent area which mainly depends on the individual experience.

Introduction

The rapid economic development requires railway to be high-speed, high-density and heavy-haul transport. It increases the likelihood of unexpected events. At present, China emergency information system lacks decision-making support function based on emergency plan management. Therefore, the emergency rescue decision-making in the field of railway transport safety management gets more attention.

At present, the domestic railway sector has done a lot of researches for railway emergency intelligent decision support system. Such as: Y. Qin¹ and others had researched the structure of the emergency management system of rail transit. Luo W.T. Luo² and others had researched the application of improved analytic hierarchy process in railway emergency plan evaluation. But these researches mainly focus on the formulation and management of contingency plans, the emergency rescue system and emergency response capability evaluation. Z.H. Zhang³ and others proposed an intelligent decision support system based on similarity. Although these methods give the rescue decision-making methods, they all have the problem of inefficient inference or inability to work in the high-dimensional data space with the large increase of data. The emergency decision-making process which is affected by many factors, such as man, car, line, environment, and the system is complicated. At the same time, with the increasing uncertainty of event information and the accumulation of cases, railway emergencies require higher efficiency decision support system. In this paper we provide a new clustering SVD algorithm based emergency decision-making reasoning approach, which is under the railway emergency decision-making researching conditions and based on the historical data analysis. The approach can not only draw lessons from the past rescue experiences, but also don’t make the system calculate slowly. It provides a rapid and scientific decision-making method for railway emergency.

Clustering SVD Algorithm

Singular Value Decomposition (SVD) is an effective mathematics feature description method,
which can extract the features of high-dimensional data and reduce the dimension, it has been widely used in image compression, signal processing, pattern analysis, etc.\[4-5\].

**Singular Value Decomposition Theory**

1. **Definition:** Let matrix $A_{m\times n} \in R^{m\times n}$, there exists unique orthogonal matrix $S_{m\times m} \in R^{m\times m}$, $D_{n\times n} \in R^{n\times n}$, and diagonal matrix $\Delta \in R^{m\times n}$, make:

$$A_{m\times n} = S_{m\times m} V_{m\times n} D_{n\times n}$$

Among them $\delta_i, i = 1, 2, \cdots, l$ is the singular value of the matrix $A_{m\times n}$, it satisfies $\delta_1 > \delta_2 > \cdots > \delta_l > 0$.

2. **Rationality of dimension reduction:** The information of diagonal matrix $V_{m\times n}$ is mainly determined by $\delta_i, i = 1, 2, \cdots, l$, the greater $\delta_i$ impact the greater diagonal matrix $V_{m\times n}$. Therefore, it has greater impact on the $A_{m\times n}$. The existing researches showed that if we delete the smaller singular value, then we should delete $S_{m\times m}$ and $D_{n\times n}$, correspondingly. After that we can obtain an approximate matrix $A'_{m\times n}$ which retains the main information of the original matrix $A_{m\times n}$. That is, assuming that $\delta_1 > \delta_2 > \cdots > \delta_t > \cdots > \delta_l > 0$ is the first $t (t \leq l)$ singular values of $A_{m\times n}$, then

$$A_{m\times n} = S_{m\times m} V_{m\times n} D_{n\times n} = A'_{m\times n} = S_{m\times m} V_{m\times n} D'_{n\times n}$$

**Clustering Analysis Principle**

Although singular value decomposition can find potential demand in high-dimensional data, the increase of data will lead to low efficiency. Clustering can divide the data into different clusters by multiple iterations, for making the similarity among the same data clusters as large as possible, and vice versa, so as to improve the system operation speed. For the specific steps of K-means clustering, see literature[4]. Cosine formula is used to calculate the similarity between vectors in this paper, as shown in Eq. (3).

$$\text{sim}(C_i, C_j) = \cos(C_i, C_j) = \frac{\sum_{k=1}^{t} C_{ik} C_{jk}}{\sqrt{\sum_{k=1}^{t} (C_{ik})^2} \sqrt{\sum_{k=1}^{t} (C_{jk})^2}}.$$  (3)

Among them, $C_i$ and $C_j$ are two column vectors with $t$ elements.

**Case Inference Model Based on Cluster Analysis**

Previous studies have shown that, orthogonal matrix $S_{m\times m}$ in $A_{m\times n} = S_{m\times m} V_{m\times n} D_{n\times n}$ mainly affects row information of $A_{m\times n}$, while $D'_{n\times n}$ mainly affects column information of $A_{m\times n}$. Therefore, this model only needs to deal with $V_{m\times n}$ and $D'_{n\times n}$ decomposed from SVD.

**Establishment of Emergency Case Inference Model**

1. **Create a Case Attribute Matrix**

In order to demonstrate the method of building a case base, this paper selected some representative cases from the railway emergency rescue cases of a certain railway bureau. Using keyword extraction method to obtain case information, see the published literature [6-7]. At the same time, according to the established situation information, as the following types of cases to obtain the value of case attributes can be obtained from the following divisions.
1. Grade attribute division: The value of number represents decision attribute, such as the occurred time and the type of the emergency. 1 is day time, 2 is night and 0 is no record time when the emergency occur.

2. Symbol attribute division: It is represented by Boolean functions, such as communication and power supply status.

3. Quantitative attribute division: It is represented by qualitative categories. When dealing with attributes (like train damage situation) qualitatively, this paper brings in the cloud model\(^5\) to generate a comprehensive evaluation cloud from a number of experts' evaluations so as to qualify the quantitative attributes. The qualitative category value can be regarded as case attribute element accordingly.

According to the above method, the case information can be converted into the data in Table 1. Thus, attribute matrix can be written as \(A = [a_{ij}]_{mn}\). The row vector represents the attribute index, the column vector represents the case information, and \(a_{ij}\) represents the value of the attribute \(i\) of the case \(j\). Based on previous data, the following statistics are available:

<table>
<thead>
<tr>
<th>Event properties</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time of accident</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Casualties</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Weather conditions</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>The event type</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Train attributes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of train</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Damage condition</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Communication status</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Power supply status</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Line properties</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line segment</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Terrain attributes</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Damage condition</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

**Case Data Processing**

(1) Perform matrix SVD decomposition

The case attribute matrix SVD is decomposed according to formula (1) to find the case attribute matrix \(S\), the event matrix \(D'\) and the diagonal matrix \(V\).

(2) Determine the intercept point of singular value

From the singular value theory, if \(\delta_i \leq \min\{m,n\}\) is taken as the intercept point, then \(rank(V) = t\) is obtained, and the diagonal matrix is changed into \(V_{\text{row}}\). Rows and columns of case attribute matrix \(S\) and case event matrix \(D'\) are all changed. From Eq. (2), the following formulation can be obtained.

\[
A_{\text{row}} = S_{\text{row}} V_{\text{row}} D'_{\text{row}}.
\]

In fact, the larger the value of \(T\) is, the weaker the effect of data dimensionality reduction of the case becomes, making it more difficult to achieve the purpose of preserving the main components of singular value decomposition. On the contrary, the role of data dimensionality reduction is more obvious. However, this may lead to incomplete feature information for lacking in a large number of information. In this paper, the intercept point is the change rate of singular value with largest change. The calculation formula of the change rate of singular value \(\varphi\) is showed in Eq. (5).

\[
\varphi = \frac{\delta_i - \delta_{i-1}}{\delta_{i-1}}, i = 1, 2, \ldots, t - 1.
\]
Cluster and Event Similarity Calculation
First, the K-means clustering is used to the reduced-dimension matrix $D'$, on the basis of selecting the number of clustering center points reasonably. Secondly, as the case dimension reduction has reduced dimensionality, the high dimension eigenvector of the emergency event must be projected as the query vector $Q_e$ of the same dimension, which is

$$Q_e = XSV^{-1}.$$  \hfill (6)

Finally, we calculate the cosine similarity between $Q_e$ and the central point of each cluster according to Eq. (3). Then we should calculate the cosine similarity between $Q_e$ and each vector in the cluster with highest similarity, and sort then according to the similarity.

Emergency Operation of Railway Emergency Decision-Making Reasoning
When a railway accident happens, the staff sends the emergency report to the railway emergency rescue decision-making support system through the private network. Through analyzing the accident information, we can get the corresponding eigenvector.

The clustering-SVD-based emergency decision-making reasoning is as follows: Choosing the most similar case in the emergency data base by the calculation of decision-making support system based on a reduction dimensionality vector of emergency attributes. Then, the successful rescue method of this case is taken as a sub-plan of the railway emergency rescue decision-making a reference for the rescue command department.

Examples of Analysis
In this section, the model is tested by a instance of emergency rescue from a certain railway Bureau. The process is as follows.

Case Description
Incident time: 12:30 at noon, February 15st, 2011.
Site situation: A passenger train got fire at the entrance of a tunnel. The locomotive ran off the rails and the number 1,2,4,5,6 cars after the locomotive got derailed. The track was destroyed for 45metres long. Heavy rain.
Casualties: 1 person seriously injured, 9 people slightly injured.

Decision Reasoning
Case Base Singular Value Decomposition
According to the method in Section 2.1.2, using the singular value decomposition to perform the case-base attribute data matrix. We obtain the case attribute matrix $S$, the diagonal matrix $V$ and the event matrix $D$.

$$\begin{pmatrix}
-0.4806 & 0.1995 & -0.1512 & 0.2905 & 0.4525 & 0.4612 & 0.1325 & -0.2666 & -0.0910 & -0.3033 & 0.1211 \\
-0.0919 & -0.0570 & -0.2822 & 0.2811 & -0.2748 & 0.1370 & -0.1027 & 0.2669 & 0.5039 & -0.3303 & -0.5319 \\
-0.2242 & -0.1449 & -0.0196 & 0.0537 & 0.0799 & -0.7376 & 0.4305 & -0.1600 & 0.1509 & -0.3695 & 0.0700 \\
-0.1567 & -0.1392 & -0.7168 & 0.2128 & -0.2599 & -0.1650 & -0.2434 & -0.2583 & -0.0980 & 0.3265 & 0.2450 \\
-0.3353 & -0.6592 & 0.1428 & -0.2732 & -0.4051 & 0.3368 & 0.2106 & -0.0835 & -0.1203 & -0.0998 & 0.0600 \\
-0.4037 & -0.1940 & -0.1991 & -0.7515 & 0.1719 & -0.0538 & -0.2196 & 0.1011 & 0.3184 & 0.0490 & 0.0189 \\
-0.3443 & -0.0946 & -0.0932 & 0.1421 & 0.1733 & -0.1234 & 0.1416 & 0.7575 & -0.3880 & 0.2191 & -0.0771 \\
-0.0278 & -0.0067 & -0.0894 & -0.0855 & 0.1520 & 0.0342 & 0.4407 & -0.3318 & 0.0019 & 0.5161 & -0.6237 \\
-0.1826 & -0.4325 & 0.2897 & 0.1436 & 0.3664 & -0.1959 & -0.6232 & -0.1882 & -0.0394 & 0.0434 & -0.2743 \\
-0.0911 & -0.1715 & 0.2142 & 0.2655 & 0.1395 & 0.0950 & 0.1420 & 0.1184 & 0.6589 & 0.4466 & 0.3869 \\
-0.4994 & 0.4680 & 0.4205 & 0.1746 & -0.4944 & -0.1313 & -0.1094 & -0.1219 & -0.0571 & 0.1664 & -0.0977
\end{pmatrix}.$$
Firstly, sort out the situation of the emergency, and this event \( X \) can be written into \( (1,1,0,1,1,1,0,0,1,1,1) \) according to the method in 2.1.1. Based on Eq. (6), the query variable can be written as

\[
Q = (-0.2210, -0.0966, -0.0694, 0.1075, -0.1186, 0.2239)^	op.
\]

Secondly, calculate the cosine similarity between the query variable and each cluster center point in accordance with Eq. (3), and find the cluster corresponding to the center point with the largest similarity.

Finally, calculate the cosine similarity between the query variable and each vector in the cluster as Table 2.
Through the analysis of decision data, we can see that the rescue plan is similar to the case \( A_{51} \) (case 11 in the database) which can be adjusted in time according to the actual emergency situation so as to make the rescue plan match the actual situation.

**Conclusion**

This paper applied a new method based on clustering and SVD to railway emergency rescue decision-making, for it not only retains the main information of the emergency, but also improves the efficiency by avoiding checking all the knowledge in the knowledge base in reasoning calculation, especially when the case data increases crazily. Meantime, the algorithm can map the high-dimensional vector to the low-dimensional space, and grasp the main information of the case, which is helpful to quickly retrieve the desired result among the information of massive and high-dimensional cases. The feasibility and maneuverability of this method for unconventional railway emergency rescue decision-making are verified by the case.

**Acknowledgement**

This research was financially supported by Gansu Province Youth Science and Technology Fund Project (145RJYA242) and Lanzhou Jiaotong University Youth Science Foundation (2014B-044).

**References**


