Study on Petroleum Standard Attention Index Calculation based on the Entropy Weight Method

Ting Xu *, Changyi Qin, Hua Zhang, Yixin Qu and Wei Fang
CNPC Tubular Goods Research Institute, Xi’an, China

*Corresponding author e-mail: xuting@cnpc.com.cn

Abstract. In this paper, the author analyzes and delves specialty distribution, attributes, attention, and other information about China’s national standards in the petroleum industry on the basis of quantitative statistics on such standards and research on attention therefor. The Entropy Weight Method (EWM) is then applied to make comprehensive measurement from three indexes of standard details query clicks, online reading clicks and download clicks, with standard attention index and corresponding ranking obtained, providing a basis for comprehensive understanding of standardization development status and trend of the petroleum industry in China.

1. Introduction
As the technical support for economic activities and social development, standards are also fundamental systems for modernization of a country’s governance system and capacity, and important means for adjustment of industry structure, transformation of development pattern and enhancement of independent innovation ability. As large process-based enterprises, petroleum enterprises feature large varieties of technical expertise, complex organization structure, high scattered production segments and complicated management on one hand, and single terminal products and basically unchanged production processes on the other hand, making it is easy to have their management standardized and highlighting the prominent role and position of standardized methods in scientific management of petroleum enterprises. In this paper, the author collects and sorts China’s national standards in the petroleum industry issued as of 2018, makes comprehensive analysis from multiple dimensions of specialty distribution, attributes, details query clicks, online reading clicks and download clicks of those standards, draws conclusion and puts forward corresponding suggestions.

2. Statistical Analysis on Standards in the Petroleum Industry
According to the Catalogue of National Standards in the Petroleum Industry (2018), there are a total of 326 prevailing national standards in the petroleum industry, which cover 17 specialties of LNG, safety, storage & transportation, geology, tubular goods, marine engineering, measurement, construction, energy saving, development, equipment, natural gas, general purpose, geophysical exploration, instruments, oil chemicals and drilling, see Table 1 for details.
Table 1. Specialty Distribution of National Standards in the Petroleum Industry (2018)

| No. | Specialties               | Standard quantity | No. | Specialties             | Standard quantity |
|-----|---------------------------|-------------------|-----|-------------------------|-------------------|
| 1   | LNG                       | 19                | 10  | Development             | 6                 |
| 2   | Safety                    | 1                 | 11  | Equipment               | 63                |
| 3   | Storage & transportation  | 6                 | 12  | Natural gas             | 59                |
| 4   | Geology                   | 13                | 13  | General purpose         | 2                 |
| 5   | Tubular goods             | 28                | 14  | Geophysical exploration | 4                 |
| 6   | Marine engineering        | 15                | 15  | Instruments             | 7                 |
| 7   | Measurement               | 46                | 16  | Oil chemicals           | 1                 |
| 8   | Construction              | 40                | 17  | Drilling                | 10                |
| 9   | Energy saving             | 6                 | 18  |                         |                   |

Among those standards, equipment, natural gas, measurement and construction standards occupy the front rank in terms of quantity, accounting for 19.3%, 18.1%, 14.1% and 12.3% respectively, with a total proportion of 63.8%. The bottom three categories in quantity are safety, oil chemicals and general purpose standards, accounting for 0.3%, 0.3% and 0.6% respectively. Specialty distribution of those standards is shown in Figure 1.

Figure 1. Specialty Distribution of National Standards in the Petroleum Industry (2018).

Among those standards, there are 33 basic standards, 4 management standards, 198 method standards, 81 product standards, 4 safety standards and 6 other standards by attributes, see Figure 2 for statistics on standard attributes.
Figure 2. Statistics on Attributes of National Standards in the Petroleum Industry (2018)

3. Entropy Indexes of Standard Attention

3.1. Principle of Entropy Weight Method

The Entropy Weight Method (EWM) is a mathematical method to calculate a comprehensive index based on the overall consideration of information and data provided by various factors, which mainly determines the weight of each index in a scientific manner based on information and data offered. EWM can be applied to accurately reflect the information and data contained in some service system evaluation indexes, and makes it possible to accurately quantify indexes with too much information.

According to the basic principles of information theory, information is a measure of the order degree of a system, while entropy is a measure of the disorder degree of a system. Given that a system is in different states and the probability of each state is $P_i (i=1, 2, \ldots, m)$, the entropy of that system is defined as

$$e = - \sum_{i=1}^{m} P_i \ln(P_i)$$  \hspace{1cm} (1)

When $P_i = 1/m (i=1, 2, \ldots, m)$, i.e. when all the states share the same probability, the entropy has its maximum value:

$$e_{\text{max}} = \ln(m)$$ \hspace{1cm} (2)

Given that there are $m$ items to be evaluated and $n$ evaluation indexes, the original evaluation matrix $R=(r_{ij})_{m \times n}$, is formed. For a certain index, the information entropy is:

$$e_j = - \sum_{i=1}^{m} P_{ij} \ln(P_{ij})$$ \hspace{1cm} (3)

It can be seen from Formula (3) that higher entropy value of an index means smaller value variation degree, less information offered, smaller role in comprehensive evaluation, and smaller corresponding weight. Smaller entropy value of an index means that greater value variation degree, more information offered, greater role comprehensive evaluation, and greater corresponding weight. Therefore, the entropy should be used to calculate the entropy weight of each index according to the degree of value variation of each index, which will be then used for weighting all evaluation indexes to obtain more objective evaluation results. As stated, the final weight of each index can be determined according to the importance of the comprehensive index and the information provided by the index. Given that there
are \( m \) items to be evaluated and a total of \( n \) evaluation indexes, the original data matrix \( R=(r_{ij})_{m \times n} \) is obtained, with the matrix form stated as follows:

\[
R = \begin{bmatrix}
  r_{11} & r_{12} & \cdots & r_{1n} \\
  r_{21} & r_{22} & \cdots & r_{2n} \\
  \vdots & \vdots & \ddots & \vdots \\
  r_{m1} & r_{m2} & \cdots & r_{mn}
\end{bmatrix}
\]  

(4)

Where, \( r_{ij} \) is the evaluation value of the \( i \)-th item under the \( j \)-th index. Basic steps to calculate weight of each index value:

Step 1: calculate the proportion \( P_{ij} \) of the index value of the \( i \)-th item under the \( j \)-th index:

\[
P_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}}
\]  

(5)

Step 2: calculate the entropy value of the \( j \)-th index:

\[
e_j = -\frac{1}{\ln(m)} \sum_{i=1}^{m} P_{ij} \ln(P_{ij})
\]  

(6)

Step 3: calculate the entropy weight \( w_j \) of the \( j \)-th index:

\[
w_j = \frac{1-e_j}{\sum_{j=1}^{n} (1-e_j)}
\]  

(7)

Step 4: determine the comprehensive weight \( \beta_j \) of the index:

Evaluators should determine the weight of index importance as \( \alpha_j \) \((j=1, 2, \ldots, n)\) according to the actual needs and purposes, and obtain the comprehensive weight of index \( j \) considering its entropy weight \( w_j \):

\[
\beta_j = \frac{\alpha_j w_j}{\sum_{i=1}^{m} \alpha_i w_i}
\]  

(8)

When value of each candidate item on index \( j \) is exactly the same, the entropy of that index reaches the maximum value of 1, and its entropy weight is zero, which means that the index cannot provide useful information to the decision maker. In this case, all candidate items under such index make no difference to the decision maker, and such index may be deleted. Therefore, entropy weight itself is the degree of discrimination to evaluation objects under the index, rather than a coefficient indicating the importance of the index.
3.2. Standard Attention Index

Top 10 out of 326 national standards in the petroleum industry in terms of standard details query clicks, online reading clicks and download clicks are defined based on the data from the National Standard Full-text Publication System, see Table 2.

**Table 2.** Statistics of Attention on National Standards in the Petroleum Industry (2018)

| No. | Serial Number | Standard Name                                                                 | details query clicks | online reading clicks | download clicks |
|-----|---------------|------------------------------------------------------------------------------|----------------------|-----------------------|-----------------|
| 1   | GB 18047      | Compressed natural gas as vehicle fuel                                        | 4499                 | 1834                  | 1338            |
| 2   | GB/T 9711     | Petroleum and natural gas industries-Steel pipe for pipeline transportation systems | 2841                 | 0                     | 1               |
| 3   | GB 32167      | Oil and gas pipeline integrity management specification                        | 2128                 | 3025                  | 861             |
| 4   | GB 17820      | Natural gas                                                                   | 1778                 | 2002                  | 890             |
| 5   | GB/T 23257    | Polyethylene coating for buried steel pipeline                                | 1572                 | 1                     | 520             |
| 6   | GB/T 9253.2   | Petroleum and natural gas industries—Threading, gauging and thread inspection of casing, tubing, and line pipe threads | 1045                 | 1                     | 182             |
| 7   | GB/T 21448    | Specification of cathodic protection for underground steel pipelines           | 1012                 | 0                     | 0               |
| 8   | GB/T 16805    | The pressure testing of steel pipelines for the transportation of petroleum, gas and highly volatile liquids | 897                  | 1                     | 122             |
| 9   | GB/T 21447    | Specifications for steel pipeline external corrosion control                   | 884                  | 1                     | 284             |
| 10  | GB/T 23505    | Petroleum and natural gas industries-Petroleum drilling and workover rigs     | 865                  | 1                     | 489             |

Formula (5) is applied for dimensionless processing of statistical data in Table 2 to form a normalized matrix as follows:

\[
P = \begin{bmatrix}
0.2568 & 0.2671 & 0.2855 \\
0.1621 & 0 & 0.0002 \\
0.1215 & 0.4406 & 0.1837 \\
0.1015 & 0.2916 & 0.1899 \\
0.0897 & 0.0001 & 0.1109 \\
0.0596 & 0.0001 & 0.0388 \\
0.0578 & 0 & 0 \\
0.0512 & 0.0001 & 0.026 \\
0.0505 & 0.0001 & 0.0606 \\
0.0494 & 0.0001 & 0.1043
\end{bmatrix}
\]

As there is zero in the matrix, the ln0 result is meaningless. The translation minimum method is adopted to eliminate the influence of zero. The translation formula is as follows:

\[
p'_{ij} = p_{ij} + \Delta
\]

(9)

Where, \( p'_{ij} \) is the value of the matrix normalization result after translation, and \( \Delta \) is the translation amplitude, which can be \( \Delta = 1 \times 10^{-5} \).

The entropy weight \( w_j \) of each index is calculated by applying Formula (6), as follows:

\[
w_j = [0.0923 \quad 0.6653 \quad 0.2423]
\]
The entropy weight is combined with the statistical data of various indexes of standard attention to obtain the attention index and comprehensive ranking of various standards, as shown in Figure 3.

![Figure 3. Comprehensive Ranking of Attention of National Standards in the Petroleum Industry.](image)

Figure 3 indicates that national standards with high attention are GB/T 32167, GB 18047 and GB 17820, with the attention index of over 1,500. Table 3 lists basic information of these standards. Attention indexes of other standards are all below 300.

| Serial Number | Standard Name | Specialties | Attributes |
|---------------|---------------|-------------|------------|
| GB/T 32167    | Oil and gas pipeline integrity management specification | Storage & transportation management |            |
| GB 18047      | Compressed natural gas as vehicle fuel | Natural gas product |            |
| GB 17820      | Natural gas | Natural gas product |            |

4. Conclusion
National standards in the petroleum industry cover diverse specialties but are unevenly distributed, mainly in four fields of equipment, natural gas, measurement and construction that account for over 63.8% of the total, with the remaining 36.2% by other 13 categories. Method standards rank first by attributes, accounting for 60.7%, followed by product standards with the proportion of 24.9%. All these show that the petroleum industry has great demand for technical standards. In addition, it is necessary to strengthen the efforts in formulation of basic, general purpose, safety and environmental protection standards aiming at adapting to the trend of China’s reform in standardization.

The subjectivity of endow weight methods such as Delphi Method and AHP can be effectively eliminated by introducing the entropy weight information theory and applying EWM to calculate the weight of standard attention in full consideration of contribution of information entropy contained in index to its role in comprehensive evaluation. Simple approach and reasonable results make it the optimal solution to objectively and truly reflect comprehensive contribution of each evaluation index, thus providing basis for multi-index comprehensive evaluation.
Acknowledgments

This work was supported by the National Key Research and Development Program of China (2019YFF0217500-2019YFF0217505).

This work was supported by Research and Development Program of China National Petroleum Corporation (No. 2019D-5005-33).

References

[1] J. F. Chen, Q. M. Wu, Y. Zhang, H. F. Nie, Thinking and practice of standard classification of petroleum industry, China Standardization. 11 (2018) 80-82.

[2] J. W. Gao, S. H. Cui, S. Cao, W. Shi, Application Status and Developing Recommendations for Standards of Petroleum Tubular Goods, Petroleum Tubular Goods & Instruments. 2 (2018) 1-8.

[3] H. J. Su, R. C. Zheng, Y. Wang, J. B. Liu, Statistical analysis on current status of work safety standards in China, Journal of Safety Science and Technology. 10 (2019) 184-188.

[4] X. Cheng, Y. Wang, Research on enterprise technical standards based on literature data statistics, Journal of The Party School of C.P.C.Qingdao Municipal Committee. 6 (2019) 73-77.

[5] F. X. Wang, Establishment and Application of Data Base of Petroleum and Petrochemical Management(1)-Standard Statistics Information Data Base, Journal of Sinopec Management Institute. 6 (2008) 35-38.

[6] Y. B. Liang, Y. C. Ke, Y. X. Lin, X. C. Zhuang, On Evaluating the Effect of Prevention and Controll Common Quality Problems in Transmission Line Engineering Based on Entropy Weight Method, Journal of Heihe University. 1 (2020) 51-59.

[7] H. P. Shen, S. Gu, F. L. Lu, X. Wu, Decision-Making Strategy for Substation Information Acquisition Based on Entropy Weight Method, Power & Energy. 12 (2016) 806-812.