Similarity Measures between Connection Numbers of Set Pair Analysis

Junjie Yang, Jianzhong Zhou, Li Liu, Yinghai Li, and Zhengjia Wu

School of Hydropower and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China
yangjunjie1998@yahoo.com.cn

Abstract. The Set Pair Analysis (SPA) is a new system analysis approach and uncertainty theory. The similarity measure between connection numbers is the key to applications of SPA in multi-attribute decision-making, pattern recognition, artificial intelligent. However, it is difficult to accurately depict similarity degree between connection numbers. The distance between connection numbers, a group of checking criterions and the similarity degree functions of connection numbers in SPA are presented in this paper to measure the similarity between connection numbers, and the rationality of such measurement is also explained by the well-designed criterions. The result shows the effectiveness of the proposed similarity measures.

Keywords: Set Pair Analysis, Similarity measures, Similarity degree function.

1 Introduction

In the real world, there are all kind of uncertainties such as fuzzy uncertainty, random uncertainty, indeterminate-known, unknown and unexpected incident uncertainty, and uncertainty which resulted from imperfective information [1]. The most successful approach to understand and manipulate the uncertainty knowledge is the fuzzy set theory proposed by Zadeh. Set Pair Analysis (SPA) theory provides another way to expressing and processing the uncertainties. The theory overlaps with many other uncertainty theory, especially with fuzzy set theory, evidence theory, Boolean reasoning methods, and rough set theory. SPA theory emphasizes the relativity and fuzziness in information processing, can identify relatively certainty information and relatively uncertainty information from the researched system. The connection number theory, that includes abundant contents and has significant meaning in the development history of mathematics, has been set up. SPA considers the connection number as a kind of number that can depict uncertain quantity, and thinks that the connection number is different from constant, variable, and super uncertain quantity essentially [23].

The similarity measure between connection numbers is the key to applications of SPA in multi-attribute decision-making, pattern recognition, artificial intelligent. However, because the connection number contains identity, discrepancy
between connection numbers, it is difficult to accurately depict similarity
degree between connection numbers. In this paper, the new similarity measures
are proposed, which are presented in Section 2 and Section 3. Finally, conclusions
are obtained in Section 4.

2 The Distance between Connection Numbers

In this section, we present the similarity measures between connection numbers
by adopting an extension of distance in functional analysis.

Definition 2.1. Let \( \mu_1 \) and \( \mu_2 \) be two connection numbers, where \( \mu_1 = a_1 + b_1i + c_1j \), \( \mu_2 = a_2 + b_2i + c_2j \). Weighted Minkowski distance between \( \mu_1 \) and \( \mu_2 \)
can be defined as follows:

\[
d_q(\mu_1, \mu_2) = \sqrt{\omega_a(a_1 - a_2)^q + \omega_b(b_1 - b_2)^q + \omega_c(c_1 - c_2)^q},
\]

where, \( \omega_a, \omega_b \) and \( \omega_c \) are weight. There are three forms of distance are as follows:

1. Hamming distance \((q=1)\)

\[
d_1(\mu_1, \mu_2) = \omega_a|a_1 - a_2| + \omega_b|b_1 - b_2| + \omega_c|c_1 - c_2|. \]

2. Hamming distance \((q=2)\)

\[
d_2(\mu_1, \mu_2) = \sqrt{\omega_a(a_1 - a_2)^2 + \omega_b(b_1 - b_2)^2 + \omega_c(c_1 - c_2)^2}. \]

3. Chebyshev distance \((q \to \infty)\)

\[
d_\infty(\mu_1, \mu_2) = \max(\omega_a|a_1 - a_2|, \omega_b|b_1 - b_2|, \omega_c|c_1 - c_2|). \]

3 Similarity Measures between Connection Numbers

In this section, a group of rationality checking criterion of similarity measures
is presented, and then similarity measures between connection numbers are pro-
posed by employing the idea of similarity degree function \[\[15\].

3.1 Checking Criterion

Let \( \mu_1, \mu_2 \) and \( \mu_3 \) be three connection numbers, where \( \mu_1 = a_1 + b_1i + c_1j \),
\( \mu_2 = a_2 + b_2i + c_2j \) and \( \mu_3 = a_3 + b_3i + c_3j \), \( \rho(\mu_1, \mu_2) \), \( \rho(\mu_1, \mu_3) \) and \( \rho(\mu_2, \mu_3) \)
denote the similarity degree function between \( \mu_1 \) and \( \mu_2, \mu_1 \) and \( \mu_3, \mu_2 \) and \( \mu_3 \)
respectively. Similarity degree function must satisfy the following criterions:

Criterion 3.1: \( 0 \leq \rho(\mu_1, \mu_2) \leq 1 \).

Criterion 3.2: (monotonicity criterion) \( \rho(\mu_1, \mu_3) \leq \min(\rho(\mu_1, \mu_2), \rho(\mu_2, \mu_3)) \),
if \( \mu_1 \leq \mu_2 \leq \mu_3 \).

Criterion 3.3: (symmetry criterion) \( \rho(\mu_1, \mu_2) = \rho(\mu_2, \mu_1) \).

Criterion 3.4: \( \rho(\mu_1, \mu_2) = 0 \) if and only if \( \mu_1 = 1+0i+0j \) and \( \mu_2 = 0+0i+1j \);
\( \rho(\mu_1, \mu_2) = 1 \) if and only if \( \mu_1 = \mu_2 \), that is \( a_1 = a_2, c_1 = c_2 \).

Criterion 3.5: \( \rho(\mu_1, \mu_2) = \rho(\mu_1', \mu_2') \), where \( \mu_1' = c + bi + aj \) is called as
complement connection number of \( \mu_1 = a + bi + cj \).