1. 结构脆弱性指标的权重确定方法——熵权法

1.1基本原理：

熵作为热力学中分子运动无序程度的度量，原本用来反映系统蕴含能量的大小。随后被引入到各领域以表征系统无序程度。当分子均匀分布在整个空间内时，分析运动表现出高度的无序性特征，系统的熵很大。当分子集中于系统中的某一子空间内时，分子运动无序程度很小，系统的熵很小。

在信息论中，熵是对不确定性的一种度量。不确定性越大，熵就越大，包含的信息量越小；不确定性越小，熵就越小，包含的信息量就越大。

根据熵的特性，可以通过计算熵值来判断一个事件的随机性及无序程度，也可以用熵值来判断某个指标的离散程度，指标的离散程度越大，该指标对综合评价的影响（权重）越大。比如样本数据在某指标下取值都相等，则该指标对总体评价的影响为0，权值为0.

熵权法是一种客观赋权法，因为它仅依赖于数据本身的离散性。

在本文中，电网拓扑结构的脆弱程度是通过结构脆弱性指标来定量描述的，不同节点对系统脆弱性的影响不同，熵是反映系统内分子离散程度的物理量，所以使用熵的概念定量描述评价指标的离散程度，即电网结构的不均衡度。进而通过熵值进行指标间的权重分配，对电网结构脆弱程度进行综合评价，并分析电力系统的结构异质性和主导因素。

在结构脆弱性指标集中，每个评价指标是考虑不同节点对电网拓扑结构的影响程度而提出的，所以当指标的熵值越大，说明各节点在该指标值下的离散程度比较小，对系统的影响程度比较均衡，系统的异质性较小；反之，熵值越小，该指标的离散程度越大，各节点的脆弱性差异也就比较大，系统具有较强的异质性。由于结构脆弱性各指标是从不同方面考虑提出的，不存在信息重合的情况。所以在评价系统结构脆弱性指标时，基于熵权法的电网结构脆弱性特征分析与评价方法是比较合理的。

1.2熵权法步骤



首先，假设有M个评价对象，根据综合评估问题的需要，选取N个评价指标，计算各评价对象的各指标数值的大小，假设第m个评价对象的第n个评价指标数值为。

然后，由于不同的评价指标在计算结果上存在量纲以及取值范围的差异，所以需要对每个评价指标的数值进行标准归一化处理。其归一化的数值用表示。（其方法为离差标准化）

计算各评价指标的熵，表达式如下：





 表示第n组随机实验中第m个随机事件发生的概率（我们假设一个节点对应的指标值越大，其受到扰动或破坏的概率越大）； 表示第n个指标的熵。

计算各评价指标的熵权，表达式如下：

（改进）

具体映射：在电力系统结构脆弱性评估中，评价对象有39个，即1-39节点，评价指标有3个，分别为电气度、电气介数和pagerank值，将每个评价指标的数值在评价指标数值总和下所占的比例作为其所发生的概率。

对熵权法的改进：

1.在标准归一化后，在指标数据里不可避免会出现数值为零的指标，而熵权法赋权的要求指标数据必须全部大于零，否则在取对数时，会出现数据异常。为了保证数据的完整性和评价的可靠性，需要对指标数据进行变换，进而对熵权法进行改进。

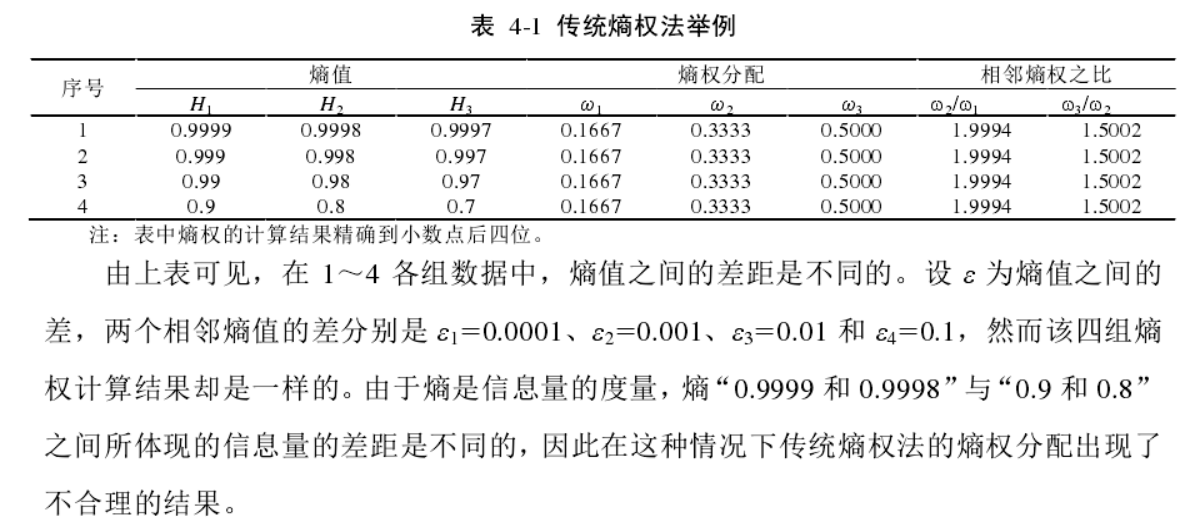
为此，为了不改变指标数据的原始分布，我们对指标数据归一化到[0.002,0.998]之间。

用均方误差来定义信息损失容忍度：

当满足，称 为在信息损失容忍度为的指标数据。（均方误差表征了变量与均值的偏离程度，用处理过的信息的均方误差与原始信息进行对比，可以得到信息损失的情况，若满足一定的容忍度，那么处理过的信息可用于系统评价）

经过计算，归一化数据的信息损失容忍度，在信息处理过程中信息损失极小。

2.改进2



不合理的地方：如第一行，三个评价指标的熵值相差不大，理论上讲，其权重分配比较均衡，按照传统熵权法公式计算，其熵权分配的差距比较大，当熵值差以数量级变动时，其权值不变。

定义：为熵值比重，其表达式为，重新定义评价指标的熵权表达式：



传统熵权法的不足之处在于，先计算了理论最大熵值与实际熵值的差，从而导致在熵值差在成数量级变化的情况下，其权值分配不变；改进的熵权法，通过先计算每个指标的熵值占总和的比例，通过重新定义指标间的熵值比重来进行权重分配，从熵权分配之初就避免了传统方法的不足，使熵权分配更加合理。



1.3实验结果

为了缩小极端值对综合评价的影响，并保证赋值的客观性，可以对极端值或者负值指标进行平移，然后再按熵权法的步骤进行计算。变换公式为：



其中， ，

在本文的指标集中，仅有数据为零的异常数据，因此只需要对零数据进行变换即可。