

## **Rough-Set-and-Genetic-Algorithm based data mining and Rule Quality Measure to hypothesize distance protective relay operation characteristics from relay event report**

### ABSTRACT

Protective relay performance analysis is only feasible by first formulating the hypothesis of expected relay operations beforehand. Traditionally, the process involved in discovering the relay operation characteristics is bogged down by the issues of differing knowledge of protection experts, meticulous manual understanding of complex relay event report and the need to have supplementary data from diverse intelligent electronic devices. This paper investigates the implementation of a novel data mining approach of integrated-Rough-Set-and-Genetic-Algorithm based rule discovery and Rule Quality Measure to hypothesize expected relay behavior in the form of an association rule from digital protective relay's resident event report. Firstly, the data mining approach of the integrated-Rough-Set-and-Genetic-Algorithm is used to discover the relay CD-decision algorithm. Subsequently, the Rule Quality Measure, combined with rule interestingness and importance judgment, deduces the relay CD-decision algorithm to the desired relay CD-association rule. The relay CD-association rule in its singularity form essentially describes the logical pattern of the correlating descriptions of conditions (i.e., attribute set C for various multifunctional protection elements) and the decision class (i.e., attribute D for trip assertion status). Using the area under the ROC curve measurements, the CD-decision algorithm has been verified to be able to predict as well as discriminate future unknown-trip-state relay events in unsupervised learning. This evaluation is necessary to allow the eventual deduction of the single relay CD-association rule to take place. The discovered CD-association rule, and thus the desired hypothesis, has been proven to be an exact manifestation of the relay operation characteristics hidden in the event report.

**Keyword:** Distance protection; Digital protective relay; Data mining; Knowledge discovery in database; Association rule; Rough set theory