



## **Understanding the rural population migration pattern of Uttarakhand using Geophysical, Geological and Socio-Economical BigData**

Kausik Chattopadhyay (1) and Pallavi Chattopadhyay (2)

(1) Institute Computer Centre, Indian Institute of Technology Roorkee, Roorkee, INDIA,(cpallavi.fes@iitr.ac.in), (2) Department of Earth Sciences, Indian Institute of Technology Roorkee, Roorkee, INDIA,(ckausik@iitr.ac.in)

Uttarakhand, a Himalayan state of India is facing a worst scenario of rural population migration for the past few decades from hill regions to the planes. While urbanization is believed to be one of the major factors for migration, how geo scientific parameters can impact the population to redraw the demographics of the hills is studied in this research. An attempt is made using density based clustering and Apriori association rule mining on 45 derived variables with a time series of 30 years to understand the rural population migration pattern. Both zone identification and origin-destination pair extraction are formulated as spatial-temporal point clustering problem and DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is applied to solve them. Specifically the population migration is formulated as a 4D point clustering problem and the relative distance between two origin – destination pair with a preference factor is used to fine tune the cluster length. In Apriori, threshold values for confidence and J-measure are kept same as for rule extraction. Rules with maximum confidence level and J-measure are obtained for an antecedent window of 18 months, consequent window of 4 months and time lag of 2 months. From the rules extracted, it can be demonstrated that almost all the geoscience indices are occurring as antecedents for migration episodes. The result demonstrates that the three districts that have registered the highest migration rates are also the districts that have witnessed maximum depletion in water sources. Even though some districts have higher number of landslide incidents, their out migration is less compared to other hill districts. However districts experiencing higher number of earthquakes are experiencing higher out migration. Upper hill region with higher precipitation experience higher migration compared to their lower hill counterpart. However this is not true when compared to the counter parts in the plane regions. Even though temperature fluctuation results in seasonal out migration, it does not have any long term impact. Resource and logistical constraints limit the frequency and extent of observations, necessitating the development of a systematic computational framework that objectively represents environmental variability at the desired spatial scale and such comprehensive big data model can be instrumental in arresting the rural migration which has been posing major threat to the livelihood of this Himalayan state.