IPED2X: A Robust Pedigree Reconstruction Algorithm for Complicated Pedigrees

Dan He 1 and Eleazar Eskin 2

1 IBM T.J. Watson Research, USA

2 Department of Computer Science, University of California, USA

Abstract

Reconstruction of family trees, or pedigree reconstruction, for a group of individuals is a fundamental problem in genetics. Some recent methods have been developed to reconstruct pedigrees using genotype data only. These methods are accurate and efficient for simple pedigrees which contain only siblings, where two individuals share the same pair of parents. A most recent method IPED2 is able to handle complicated pedigrees with half-sibling relationships, where two individuals share only one parent. However, the method is shown to miss many true positive half-sibling relationships as it removes all suspicious half-sibling relationships during the parent construction process. In this work, we propose a novel method IPED2X, which deploys a more robust algorithm for parent construction in the pedigrees by considering more possible operations rather than simple deletion. We convert the parent construction problem into a graph labeling problem and propose a more e_ective labeling algorithm. We show in our experiments that IPED2X is more powerful on capturing the true half-sibling relationships, which further leads to better reconstruction accuracy.