SUMMARY OF Ph.D. DISSERTATION

School Fundamental Science and Technology	Student Identification Number	SURNAME, First name UEHARA, Hiroaki

Title

Algorithms for the computation of approximate
posterior probabilities based on group testing

Abstract

In 1940s, Dorfman showed that group testing is quite effective for screening positive items from a large number of items. In a group test, multiple items which have the same background are allocated into several groups, so that positiveness of each items can be inferred from the result of the group test. Group testing has been extensively used in various fields of computer science, life science, and so on. A recent significant application of group testing can be found in DNA library screening, but a further development of better algorithm for computing the posterior probability of the positiveness of each item is needed. In this thesis, two new algorithms are reported for computing the posterior probability, based on the Markov chain pool result decoder (MCPD) proposed by Knill et al. (1996) and compared with MCPD by means of simulations.

In Chapter 1, a brief history for group testing is given and a probabilistic model for group testing is described.

In Chapter 2, an algorithm called Bayesian network pool result decoder (BNPD) is proposed for computing the posterior probability, which is based on the belief propagation algorithm of Bayesian network. Another algorithm proposed is an algorithm called CCCP pool result decoder (CCPD), which is based on the concave-convex procedure (CCCP). It has been developed to cope with the constraint in BNPD that the Tanner graph which indicates allocation of items into each group must have no cycle of length four. The bias of the computed posterior probability by BNPD and CCPD is also theoretically evaluated. It is showed that most significant part of the bias is zero as far as the Tanner graph has no cycle of length four but otherwise remains.

In Chapter 3, the performance of BNPD and CCPD is compared, by simulation, with that of MCPD. Efficiency of CCPD is compared with that of MCPD after showing advantages of BNPD when the Tanner graph has no cycles of length four. It is also examined that allocation of the items according to a regular Tanner graph which has no cycles of length four shows a better performance than that according to graphs generated randomly.

In Chapter 4, we conclude that it is important to allocate items into groups according to a Tanner graph with no cycle of length four for the use of BNPD, CCPD or MCPD. Otherwise, it may cause a significant bias of the computed posterior probability.