Applied Soft Computing 52 (2017) 1–13

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

# An adaptive ant colony optimization algorithm for constructing cognitive diagnosis tests☆

Ying Lin[a], Yue-Jiao Gong[b], Jun Zhang[b,*]

[a] Department of Psychology, Sun Yat-Sen University, Guangzhou, China
[b] School of Computer Science & Engineering, South China University of Technology, Guangzhou, China

ARTICLE INFO

*Article history:*
Received 26 July 2015
Received in revised form
17 September 2016
Accepted 24 November 2016
Available online 14 December 2016

*Keywords:*
Ant colony optimization (ACO)
Cognitive diagnosis model (CDM)
Test construction

ABSTRACT

A critical issue in the applications of cognitive diagnosis models (CDMs) is how to construct a feasible test that achieves the optimal statistical performance for a given purpose. As it is hard to mathematically formulate the statistical performance of a CDM test based on the items used, exact algorithms are inapplicable to the problem. Existing test construction heuristics, however, suffer from either limited applicability or slow convergence. In order to efficiently approximate the optimal CDM test for different construction purposes, this paper proposes a novel test construction method based on ant colony optimization (ACO-TC). This method guides the test construction procedure with pheromone that represents previous construction experience and heuristic information that combines different item discrimination indices. Each test constructed is evaluated through simulation to ensure convergence towards the actual optimum. To further improve the search efficiency, an adaptation strategy is developed, which adjusts the design of heuristic information automatically according to the problem instance and the search stage. The effectiveness and efficiency of the proposed method is validated through a series of experiments with different conditions. Results show that compared with traditional test construction methods of CDMs, the proposed ACO-TC method can find a test with better statistical performance at a faster speed.


## 1. Introduction

Assessment of subjects' latent traits has long been an important and challenging topic in psychometry. In classic item response theory (IRT), a subject's latent trait is measured by a total score calculated based on his/her responses to assessment items. However, a score can only provide limited information. Both subjects and practitioners are desired for methods that can give more detailed feedback through response analyses. Driven by the demand, a new class of psychometric models named as cognitive diagnosis models (CDMs) were developed. CDMs were first used in education to assess a student's learning progress on a subject by estimating his/her mastery level on each fine-grained skill (often referred to as an *attribute*) [1,2]. With CDMs diagnosing the student's strengths and weakness at the attribute level, individualized instructions and targeted remediation can be developed accordingly. Encouraged by the success in education, CDMs have been extended to other fields of psychometry, including assessment of broad cognitive skills [3,4] and psychological disorders like pathological gambling [5].

The effectiveness of CDMs as psychometric models has been admitted by more and more researchers. However, in contrast to the fast development of CDMs in theories, their real-world applications progress at a slow pace. A main obstacle lies in test construction. That is, given a calibrated item bank, how to select items so that the resulting test achieves a satisfying statistical performance while fulfilling all the practical constraints. A primary criterion for evaluating the statistical performance of a CDM test is the diagnostic accuracy. However, there are no mathematical expressions that can formulate the relationship between the diagnostic accuracy of a test and the choice of test items. It is thus difficult to design or apply exact algorithms to the test construction problem of CDMs. Heuristic methods that can approximate the optimum in reasonable time are more promising. In the literature, depending on the way to evaluate test quality, the heuristics for constructing a CDM test can be classified into two categories: index-oriented and simulation-oriented.

The index-oriented methods were developed based on *item discrimination indices*. In CDMs, each item has two types of discrimination indices: cognitive diagnosis information (CDI) [6] and attribute-level discrimination indices (ADI) [7]. CDI measures the

☆ This work was supported in part by NSFC grant no. 61309003 and in part by the NSFC Key Project no. 61332002.
* Corresponding author.
*E-mail address:* junzhang@ieee.org (J. Zhang).

http://dx.doi.org/10.1016/j.asoc.2016.11.042
1568-4946/

ability of an item to distinguish different attribute mastery patterns. ADI measure the ability of an item to distinguish different mastery levels on each attribute. A larger index value implies a better distinguish ability. Therefore, a test that has a larger sum of item discrimination indices should offer a better diagnostic accuracy in the corresponding perspective. Based on the idea, Hensons and Douglas [6] used the items with the largest CDI to compose a test that achieves high overall diagnostic accuracy. Finkelman et al. [8] applied the binary programming technique to maximize the minimum sum of ADI among all the attributes, resulting in a test that achieves high diagnostic accuracy at the attribute level. Kuo et al. [9] modified the CDI and ADI so that the hierarchy of attributes and the ratio of test length to the number of attributes can be taken into account for test construction. Simulations showed that the above methods can generate better tests than random selection in the sense of achieving high diagnostic accuracy. However, they suffer setbacks in the following two aspects. First, the possible interactions between items are overlooked. Although the resulting test is optimal in terms of the objective function defined based on item discrimination indices, the diagnostic accuracy may still be suboptimal. Second, there are circumstances that test construction has a purpose other than maximizing the diagnostic accuracy, e.g., requiring the diagnostic accuracy to meet a targeted precision [10–12]. In that case, the effectiveness of the index-oriented methods is largely impaired.

Instead of searching for the optimal test in terms of the objective function defined based on item discrimination indices, the simulation-oriented methods search for the test that performs the best in simulation. By simulating subjects' responses to a test, the simulation-oriented methods take the possible interactions between test items into account. The diagnostic accuracy of the test can then be obtained by comparing subjects' attribute mastery levels estimated from the responses to the real levels known in advance. Relying on the above simulation-based test evaluation process, the simulation-oriented method can search for the optimal test for different construction purposes. However, devising an effective simulation-oriented method is a challenging task as no mathematical properties (e.g., gradient or derivative) can be acquired from the problem to guide the search. Finkelman et al. [13] proposed a genetic algorithm (GA) to accomplish the task. GA is a meta-heuristic that evolves a population of candidate solutions towards the optimum by imitating the evolution process of organisms [14]. In Finkelman's GA, the initial population is composed of tests generated by index-oriented methods. Then a number of new tests are produced by replacing each item in an existing test with a feasible one. After evaluating the performance of each newly generated test based on simulation, the best several tests are selected to form the population of next iteration. By iterating the above steps of reproduction, evaluation, and selection, the GA evolves the population towards the optimal test on the given construction purpose. Although Finkelman's GA offers better test quality and higher flexibility, it suffers from long computational time due to the compute-intensive evaluation step and the inefficient search paradigm. As the scale of the test construction problem increases, the disadvantage may become more prominent.

Concluded from the above, CDMs are still expecting a test construction method that can find a high-quality test efficiently for different construction purposes. For filling the gap, this paper proposes a test construction method based on ant colony optimization (ACO). ACO is a meta-heuristic inspired from the ants' foraging behavior in nature [15–17]. Each artificial ant in the algorithms is a stochastic procedure that constructs a solution. Like GA, ACO algorithms can search without using mathematical properties of the problem, but instead of searching at complete random, ACO algorithms can improve the search efficiency by guiding the ants' construction behavior with search experience (represented by *pheromone*) and domain knowledge (represented by *heuristic information*). In particular, the proposed ACO-based test construction (ACO-TC) method defines heuristic information based on item discrimination indices. An adaptation strategy is developed to adjust the design of heuristic information so that it best fits the problem instance and the search stage. By doing so, the useful information behind item discrimination indices can be used to guide the ants' construction procedures, which accelerates the speed of finding a high-quality test. Meanwhile, all the tests constructed are evaluated based on their performances in simulation. The convergence towards the optimal test is thus guaranteed regardless of the construction purpose. As such, ACO-TC becomes a simulation-oriented method that uses item discrimination indices to accelerate the search speed. It is expected to achieve high efficiency and wide applicability simultaneously. For validation, a series of experiments are performed under a variety of conditions. Results are compared with classic test construction methods of CDMs to show the effectiveness and efficiency of the proposed method.

The main contribution of this paper lies in the following two aspects. First, an ACO-based test construction method is proposed for CDMs. The proposed method can be tailored to different construction purposes and achieve better performance than classic methods in terms of both test quality and search speed. As far as we know, it is the first effort of using ACO to solve the test construction problem of CDMs. Second, an adaptation strategy is developed for ACO to automatically adjust the design of heuristic information according to the problem instance and the search stage. The strategy is a general method that can be applied to different problems and different ACO variants. ACO can benefit from the strategy as it offers a more appropriate design of heuristic information and increases the diversity in solution construction behavior.

The remainder of this paper is organized as follows. As an introduction to background knowledge, Section 2 presents a brief review of CDMs and ACO. Section 3 details the implementation of the proposed ACO-TC method. The adaptation strategy of heuristic information is introduced in Section 4. The simulation results and discussions are shown in Section 5. Section 6 draws a conclusion to the whole paper.

## 2. Background

This section contains three parts. The first part reviews the essential idea of CDMs and introduces the model used for exemplifying the proposed method. The second part describes the formal definitions of the two types of item discrimination indices, CDI and ADI. In the last part, the general framework of ACO algorithms is presented.

### 2.1. Cognitive diagnosis models (CDMs)

CDMs are statistical models that define the probability of a correct response to an item based on attribute mastery patterns. Given $K$ attributes, an attribute mastery pattern is a $K$-dimensional vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_K)$, where $\alpha_k$ denotes the mastery level of attribute $k$ and $k = 1, 2, \ldots, K$. Suppose that there are $L$ mastery levels for each attribute. The total number of attribute mastery patterns is $L^K$. A cognitive diagnosis assessment aims to classify each subject into one of the $L^K$ patterns so that the likelihood of the observed responses is maximized. Based on different ideas of how attribute mastery patterns affect responses, various CDMs have been developed. For overview and taxonomy of the notable CDMs in recent years please refer to [18–20]. Since the proposed algorithm ACO-TC requires information of a concrete CDM for test quality evaluation and heuristic information computation, this paper takes the reduced reparameterized unified model (r-RUM) [21] as an