Science in China Series D: Earth Sciences
© 2007 SCIENCE IN CHINA PRESS
Springer

Discovery of transition rules for geographical cellular automata by using ant colony optimization

LIU XiaoPing[1], LI Xia[1†], Anthony Gar-On YEH[2], HE JinQiang[1] & TAO Jia[1]

[1] School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China;
[2] Center of Urban Planning and Environmental Management, The University of Hong Kong, Hong Kong SAR, China

A new intelligent algorithm of geographical cellular automata (CA) based on ant colony optimization (ACO) is proposed in this paper. CA is capable of simulating the evolution of complex geographical phenomena, and the core of CA models is how to define transition rules. However, most of the transition rules are defined by mathematical equations, and are hence not explicit. When the study area is complicated, it is much more difficult to extract parameters for geographical CA. As a result, ACO is applied to geographical CA to automatically and intelligently obtain transition rules in this paper. The transition rules extracted by ACO are defined as logical expressions rather than implicit mathematical equations to describe the complex relationships of the nature, and easy for people to understand. The ACO-CA model was applied to simulating rural-urban land conversions in Guangzhou City, China, and appropriate simulation results were generated. Compared with See5.0 decision tree model, ACO-CA is more suitable to discovering transition rules for geographical CA.

ant colony optimization, CA, geographical simulation, artificial intelligence

1 Introduction

Cellular automata (CA) was first proposed by Ulam in the 1940s and soon used by Von Neumann to investigate the logical nature of self-reproducible systems. CA has strong capabilities in simulating the tempo-spatial evolution of complex systems. One key feature of CA is that complex global spatial patterns can be generated by some simple and local rules, this 'bottom-up' approach coincides with complexity theories that a complex system comes from the interactions of some simple subsystems. CA can be used to simulate the unexpected behaviors of complex systems which cannot be represented by concrete equations. CA is suitable for simulation and prediction of complex geographical processes. At the end of the 1980s, Couceleis put forward the theoretical framework for CA's application in geosciences in details; especially the experiment of urban expansion simulation has contributed a lot to the study of urban evolution[1−3]. Recently, CA has been applied to the simulation of population dynamics[2], wildfire propagation[4], urban evolution[5−10], and land-use changes[5,11], of which it may be the most successful example for urban simulation to solve geographic problems[5,6,8]. Such researches show that CA is suitable for simulating and predicting complex geographical processes.

The definition of transition rule is the core of the CA model, as transition rule expresses the logical relationships among the simulated processes, and determines the consequence of their spatial evolution. However, the determination of transition rules is very tedious. Much more methodologies were promoted by scholars. Simple visual comparison was proposed by Clarke to obtain model parameters[8]. Possible scenarios are generated by

Received November 13 2006; accepted February 13, 2007
doi: 10.1007/s11430-007-0083-z
†Corresponding author (email: lixia@mail.sysu.edu.cn)
Supported by the National Outstanding Youth Foundation of China (Grant No. 40525002), the National 863 Project of China (2006AA12Z206), the National Natural Science Foundation of China (Grant No. 40471105), and the "985 Project" of GIS and Remote Sensing for Geosciences from the Ministry of Education of China (Grant No. 105203200400006)

exploring various combinations of parameter values, best set of which is determined with visual comparison between the simulations and actual ones. However, it is difficult to find the best option for numerous combinations based on variables. Hierarchical analysis procedure (AHP) was then used to heuristically define parameter values[10]. And logistic regression model was further presented to calibrate CA[12]. These models have been extensively applied due to their simplicity and practicability, but they also pose problems to deal with complicated geography phenomena for linear feature. Then a neural-network CA model was developed to automatically obtain parameter values for complex relationship[13]. Because of black-box character, it is difficult for users to comprehend the meanings of these parameters and the mechanisms. Decision tree model was later presented to recover transition rules of CA[14], however, it proves to be easily vulnerable to local optimization. With the development of CA model, kernel-based learning machine was induced to obtain nonlinear transition rules in high-dimension feature space[15], which is also constrained by implicit physical implications of transition rules and a large amount of calculation.

Additionally, when the study area becomes more complex, the exiting methods pose a problem to derive CA model structure and related parameters. Hence, it's necessary to introduce intelligent methods to effectively retrieve transition rules. Being a kind of artificial intelligence optimization, ant colony is in fact a complex multi-agent system, which can be used for discovering CA transition rules. ACO, initially proposed by Colorni and Dorigo[16], is a system based on agents which simulated the natural behavior of ants, including mechanisms of cooperation and adaptation. Complex tasks, such as optimized route for seeking foods, can be effectively fulfilled by the mutual cooperation between ants. During the optimization processes, each ant agent can make a random choice according to the information of the routes. Though without centralized control, the whole system can still be optimized and the shortest route can be easily located. Because such ant colony system is strongly robust, the whole solution is not easily affected by one's or several agent' failure. As a result, ACO is a typical swarm intelligence-based heuristic. This is advantageous because it allows the system to use a mechanism of positive feedback between agents as a search mechanism. Recently, ACO has now become a hot topic in artificial intelligence field[17,18]. However, applying ACO to geosciences is far less reported. Actually, the 'bottom-up' approaches adopted in ACO for achieving complex task through cooperation among agents coincides well with CA, therefore, it is appropriate for ACO to be used for extraction of CA transition rules.

This paper will propose a new method to obtain CA's transition rules based on the techniques of ant colony optimization. An Ant-Miner program will be developed for discovering transition rules. It will retrieve optimized rules through simulating the behavior of ants' seeking foods using the shortest paths. Besides, this model will be applied to the simulation of rural-urban land conversions in Guangzhou. Compared with See5.0 decision tree model, the results show that ACO method is more suitable for CA model. No such studies have been reported so far.

2 Ant colony optimization

The artificial ant colony optimization (ACO) was based on the ants' behaviors of finding the shortest path when seeking foods without use of the visual information[17]. This intriguing ability of blind ants has been extensively studied by ethologists. They discovered that, in order to exchange information about which path should be followed, ants communicate with one another by means of pheromone, a kind of secretion, which is unique to ants. While walking, ants deposit pheromone on the ground, and follow, in probability, pheromone previously deposited by other ants. It can be sensed by moving ants to direct their movement, but it evaporates with time. When the number of ants increases along a certain path, the trail and amount of the pheromones deposited will increase, resulting in a higher probability for other ants to choose this path. In this way, ants can locate the shortest path from ant nests to food sources through indirect communication among individuals. This process can be described as a loop of positive feedback, in which the probability that an ant chooses a path is proportional to the number of ants that have already passed by that path[17].

The above food-seeking process based on positive feedback information indicates that ACO is self-adaptive. The process of seeking food by an ant colony is illustrated by Figure 1. It shows that swarm intelligence can be embodied in ant colony optimization. If there are no obstacles between ant nests and food sources, the shortest path is in a straight line (Figure 1(a)). The attraction

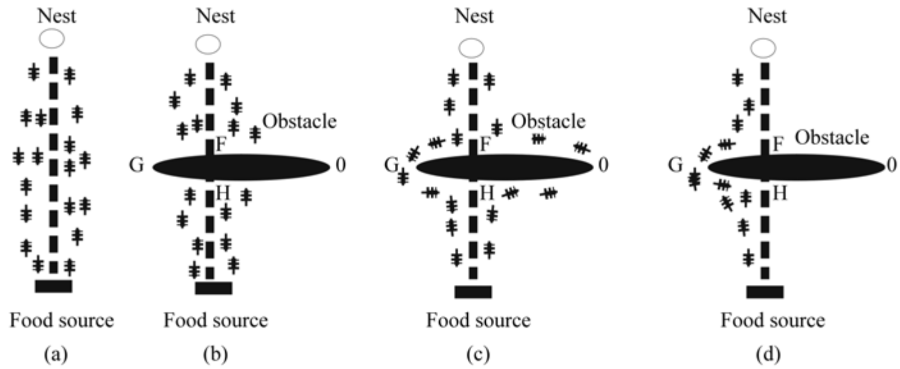

Figure 1 The path choice behaviors of ants in seeking foods.

is that an ant colony can not only fulfill complex tasks but also adapt to environmental changes. For example, if there is an obstacle occurring on the route, ants can use swarm intelligence to find the optimal solution. At the beginning, ants select various paths by identical probability (Figure 1(b)). During their movement, ants will deposit pheromone on paths that they passed by. Since the path F-G-H is shorter than F-O-H, the ants selecting the path F-G-H will reach the food source earlier than those selecting the path F-O-H. The amount of pheromone will be deposited more on H-G-F than on H-O-F. This will result in more ants to select the path H-G-F (Figure 1(c)). With the pheromones on the longer path gradually disappearing due to evaporation, all ants will move to the shortest path under this exploration process (Figure 1(d)).

3 ACO-based geographical CA

Characterized by the positive feedback process, ant colony optimization is a kind of abstraction and simulation of true ants' behaviors on searching for food[17,18]. Satisfactory results have been obtained in solving traveling salesman problem (TSP), distribution problem, data clustering, combinatorial optimization and network routing by using ant-colony optimization[19−21]. However, the studies on discovering classification rules using ACO are still at the initial stage. The method of ACO-based rule discovery was first proposed by Parpinelli[22]. The strategy of seeking foods by ant colony was applied to the extraction of optimal rules in databases. This ACO method can effectively solve nonlinear problems because of its capability of self-learning. It is especially useful for analyzing complex geographical phenomena by updating knowledge bases according to environmental changes and past behaviors. In this paper, the ant colony-based method will be modified and applied to the rule induction for discovering transition rules of CA.

Route search by an ant colony is to find the links between attribute nodes and class nodes. The attribute node can only be selected once and must be associated with a class node. As shown in Figure 2, each route corresponds to a classification rule, and data-mining for a classification rule can be regarded as searching for optimal route. A rule can be randomly generated at the start. The rule can be represented as follows:

$$\text{IF} <\text{term}_1 \text{ AND term}_2 \text{ AND}\ldots\ldots> \text{THEN} \langle\text{class}\rangle \quad (1)$$

where term_i are condition items, and the logical combination of condition items can be expressed as the triple ⟨*attribute, operator, value*⟩, *where value is a value* belonging to the domain of *attribute*. The operator element in the triple is a relation operator. The ⟨class⟩ is the prediction for this case.

It should be noted that the original continuous values must be converted into discrete ones for facilitating the route search. If the original values are V_1, V_2, …, V_n

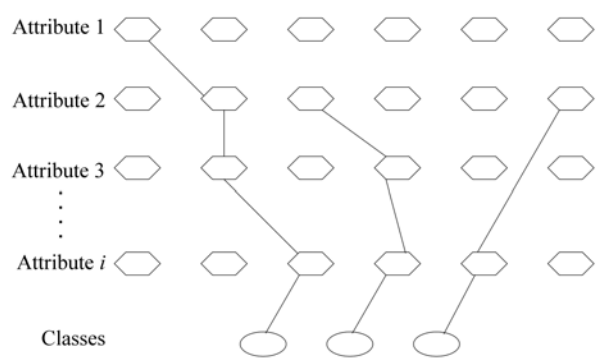

Figure 2 Route corresponding to classification rule derived from Ant-Miner.

for the attributes of A_1, A_2, …, A_n, these values should be discretized as V_{11}, V_{12}, …, V_{21}, V_{22}, …, V_{nm}. The following sections will provide the detailed procedure for applying ACO to discovery transition rules of CA.

Knowledge discovery of classification rules based on ant colony principles can be divided into three stages. First starting from an empty route, node can be selected repeatedly and added to the route until a complete route is acquired, that is, a rule is constructed; then the rule will be pruned; finally the pheromone amount on all the routes shall be updated, which will definitely affect the rule construction for the next ant.

3.1 Rule construction

Rule construction imitates the food seeking behavior of ants. The search procedure is to select nodes repeatedly until an integrated route is constructed. Theoretically, node selection can be completely random, but this will result in much long time for computation.

A heuristic function can be designed to guide ants' searching by reducing computation time. This function is critical for the computation to reach the convergence more quickly. The information entropy can be used to define this function, in which the heuristic value for each attribute node is proportional to its classification capability[22]. In this paper, a heuristic function based on the statistical attribute of the data (frequency) is designed, in which the heuristic value η_{ij} of the condition item term_{ij} is defined as[21]

$$\eta_{ij} = \frac{\max(\sum_n \mathrm{freq}T_{ij}^1, \sum_n \mathrm{freq}T_{ij}^2, \cdots \sum_n \mathrm{freq}T_{ij}^k)}{\sum_n T_{ij}}, \qquad (2)$$

where η_{ij} is denoted as the density-based heuristic value of the condition item term_{ij}, T_{ij} refers to the number of cases fitting to the condition term term_{ij}, $\mathrm{freq}T_{ij}^w$ is the frequency of class w in T_{ij}. The record that satisfies the condition part of the rule should be removed after a final rule has been obtained. Therefore, the values for $\max(\sum_n \mathrm{freq}T_{ij}^1, \sum_n \mathrm{freq}T_{ij}^2, \cdots \sum_n \mathrm{freq}T_{ij}^k)$ and $\sum_n T_{ij}$ are updated after a final rule has been found.

The other two parameters, the amount of pheromone and the probability for the attribute nodes to be selected, are also important to the generation of rules. When a route is found by an ant, the thickness of pheromone for all the nodes in this route will be initialized to the same value:

$$\tau_{ij}(t=0) = \frac{1}{\sum_{i=1}^{a} b_i}, \qquad (3)$$

where τ_{ij} is the amount of pheromone for the condition term term_{ij}, a is the sum of attributes (excluding the class attributes) in the databank, b_i refers to any possible value of attribute i.

The roulette mechanism is adopted to decide which attribute node will be selected. For each attribute row, the probability for its node term_{ij} to be selected is calculated according to the following formula:

$$P_{ij}(t) = \frac{\tau_{ij}(t)\cdot\eta_{ij}(t)}{\sum_{i=1}^{a}\sum_{j=1}^{b_i}\tau_{ij}(t)\cdot\eta_{ij}(t)}. \qquad (4)$$

The selected attribute nodes will be continuously added to the route until all attributes (including class attributes) are selected to form a final route (a classification rule). The validity of this rule can be assessed by using the following formula[22]:

$$Q = \left(\frac{\mathrm{TruePos}}{\mathrm{TruePos}+\mathrm{FalseNeg}}\right)\cdot\left(\frac{\mathrm{TrueNeg}}{\mathrm{FalsePos}+\mathrm{TrueNeg}}\right), \qquad (5)$$

where TruePos (true positives) is the number of cases covered by the rule that have the class predicted by the rule; FalsePos (false positives) is the number of cases covered by the rule that have a class different from the class predicted by the rule; FalseNeg (false negatives) is the number of cases that are not covered by the rule but that have the class predicted by the rule; TrueNeg (true negatives) is the number of cases that are not covered by the rule and that do not have the class predicted by the rule. The larger the value of Q, the higher the quality of the rule.

3.2 Rule pruning

The next step is to prune the inducted rules for better classification results. Rule pruning is a commonplace technique in data mining[23]. The goal of rule pruning is to remove irrelevant terms that might have been unduly included in the rule. Above rule induction may create a large set of rules, which are difficult to interpret. Some rules have little contribution to the classification and may even bring negative impacts on the accuracy. Moreover, selecting route nodes repeatedly may result in over-fitting of training data. Therefore, rule pruning potentially increases the predictive power of the rule, helping to avoid its overfitting to the training data. Another motivation for rule pruning is that it improves the simplicity of the rule, since a shorter rule is usually easier to be understood by the user than a longer one[22].

One simple way of rule pruning is to iteratively remove one-term-at-a-time from the rule while this process improves the quality of the rule. More details, in the first iteration one starts with the full rule. Then tried to remove each of the terms of the rule, and the quality of the resulting rule is computed by eq. (5). It should be noted that this step might involve replacing the class in the rule consequent, since the majority class in the cases covered by the pruned rule can be different from the majority class in the cases covered by the original rule. In the next iteration it is removed again the term whose removal most improves the quality of the rule, and so on. This process is repeated until the rule has just one term or until there is no term whose removal will improve the quality of the rule[22]. The pseudo-code for rule pruning is as follows:

```
No_of_terms=No_of_attributes–1;
validity_newrule=1
WHILE (validity_newrule>validity_previousrule)
validity_newrule=validity_previousrule;
   FOR j=1 TO No_of_terms
        Remove term_j from the rule
        Calculate the validity_newrule_j
             IF (validity_newrule< validity_newrule_j) THEN
                  validity_newrule= validity_newrule_j
                  Obtaining a better rule
             END IF
   NEXT j
No_of_terms = No_of_terms–1
LOOP
```

3.3 Pheromone updating

In one iteration, after the rules constructed by artificial ants are pruned, the amount of pheromone at all route nodes will be updated according to the efficiency of the rule for classification. The pheromone amount at route nodes covered by the rule will be increased, while the pheromone amount at route nodes not covered by the rule will be decreased (here the evaporation coefficient ρ will be introduced). The pheromone amount at each attribute node will be updated according to the following formula:

$$\tau_{ij}(t+1) = (1-\rho)\cdot\tau_{ij}(t) + \Delta\tau_{ij}(t), \tag{6}$$

$$\Delta\tau_{ij}(t) = \sum_k \Delta\tau_{ij}^k(t), \tag{7}$$

$$\Delta\tau_{ij}^k(t) = \begin{cases} \dfrac{Q_k}{1+Q_k} & \text{if the } k \text{ artificial ant passed over node term}_{ij} \\ 0 & \text{else} \end{cases}, \tag{8}$$

where ρ is the pheromone evaporation coefficient, Q_k is the quality of a classification rule, $\Delta\tau_{ij}^k(t)$ is the pheromone amount remained on the node term$_{ij}$ by the k ant. When the pheromone amount at all attribute nodes has been updated, the next ant will start its searching. When many ants locate continuously the same route in their search, this process is convergent. Otherwise, this process will be repeated until all ants complete their search. In this iteration process, each ant will construct a rule, but only the rule of the best quality can be preserved and be regarded as the ultimate classification rule. Other rules of poorer quality will be discarded. Iteration will be repeated until the number of remaining training classes is less than the predefined number of classes.

3.4 Discovering transition rules of CA by using ACO

Characterized by high self-learning capability, ant colony optimization gradually updates its own knowledge base corresponding to environmental changes and previous behaviors, and in this way, realizing problem solution. As a result, it can be used to effectively solve non-linear problems, and is particularly suitable for complex geographical phenomena. Additionally, the 'bottom-up' approaches adopted in ACO coincide well with CA, which is very suitable for extracting CA transition rules. In this paper, based on classification rules from Ant-Miner, transition rules are automatically derived from training data sets, and simultaneously model is calibrated during the process. The structure of ACO-based geographical CA model is shown as Figure 3.

The ACO-CA model consists of two parts: rules discovery and city development simulation. Remote sensing data in the past two years will be utilized to monitor the growth of the city, and transition rules are mainly discovered by using Ant-Miner, which is implemented through Visual Basic 6.0 programming. The pseudo-code for discovering transition rules of CA is as follows:

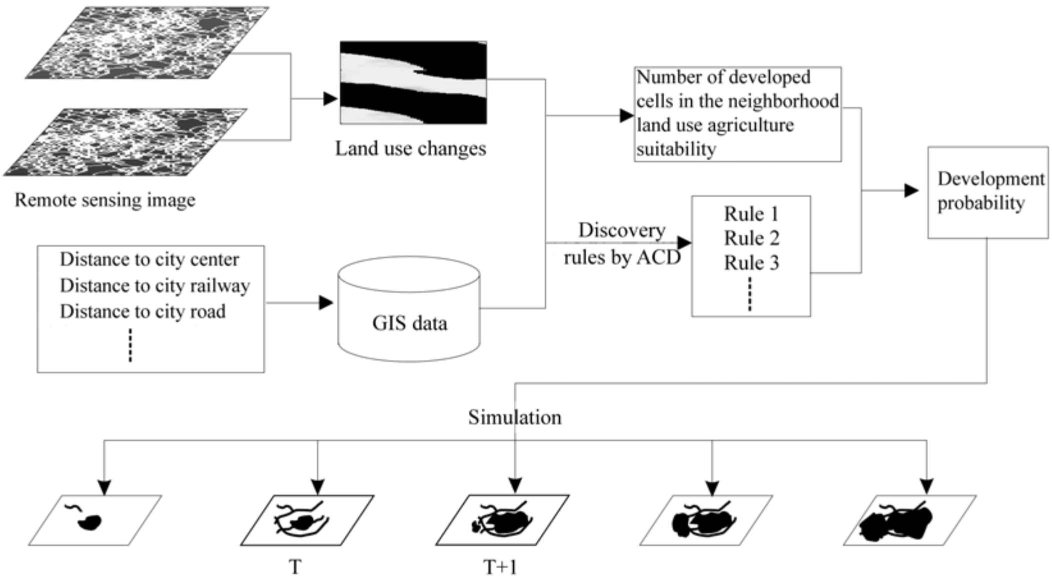

Figure 3 ACO-based approach for cellular automata.

The original trainingSet
Discretization of the original TrainingSet
DiscoveredRuleList=[] /* rule list is initialized with an emptylist */
WHILE (TrainingSet > Max_uncovered_cases)
Initialize all nodes with the same amount of pheromone
calculation the $\eta_{ij}\eta_{ij}$ of the training data for all nodes
i = 1 /* ant index */
WHILE (i<No_of_ants and m<No_rules_converg)
FOR j=1 TO No_of_attributes
Select a node of the attribute
NEXT j
Obtaining Rulei
rules pruning
IF (Rule$_i$ is equal to Rule$_{i-1}$ THEN
m=m+1
ELSE
m=1
END IF
pheromone update
i=i+1
LOOP
Select the best rule R_{best} among all rules constructed by all the ants;
Add rule R_{best} to DiscoveredRuleList
TrainingSet = TrainingSet-{set of cases covered by R_{best}}
LOOP

Ant-Miner-derived transition rules are different from equation-based CA transition rules. The latter are implicitly expressed with mathematical equations, while the former are explicitly defined. For example:

Rule 1:

IF Distance to urban centers <8 km,
Distance to trunk roads <0.5 km,
The number of developed cells in the neighborhood >5,
Land use types = farmland.

THEN Development is allowed (confidence = 0.98).

Rule 2:

IF Distance to urban centers >50 km,
Number of developed cells in the neighborhood <2, Land-use types = forestland.

THEN Development is prohibited (confidence = 0.90).

Simulation is implemented through secondary development with ArcObjects plus Visual Basic 6.0. During the process, changes of the central cell in the neighborhood are dynamically computed, and the number of urbanized cells was easily calculated with Focal function of ArcGIS. It is noted that the observation interval (ΔT) between remote sensing images is generally far greater than the iteration interval (Δt) of CA simulation. It may be ideal if the observation interval (ΔT) is equal or close to the iteration interval (Δt) so that transition rules mined can be used directly in urban simulation[14]. As a result, it is necessary to determine the amount of land use conversion in the iteration interval (Δt) in CA model. Firstly, the number of iterations (K) of CA model during the period of iteration is represented as follows:

$$K = \Delta T / \Delta t. \tag{9}$$

Secondly, amount of land use conversion (ΔQ_0) can be determined from remote sensing for the larger observation interval (ΔT). As $\Delta T > \Delta t$, only a portion of land use conversion took place in the iteration Δt. Amount of land use conversion between t and t+1 can be calculated as

$$\Delta q_0 = \Delta Q_0 / K, \tag{10}$$

where Δq_0 is the amount of land use conversion for the iteration interval Δt.

4 Application and simulation results

4.1 Test area and spatial data

Guangzhou City, in the Pearl River Delta of China was selected as test area. TM satellite images in 1988, 1993 and 2002 were used to provide actual urban areas, which were divided into observation data including TM data in 1988, 1993 for detailed transition rules and test data of 2002 for capturing the urban development trend. The probability of land development is related to a series of spatial distance variables, neighborhood conditions, and physical attributes[5,24], which are derived from remote sensing and GIS data, are used for the discovery rules, listed in Table 1.

To obtain transition rules, stratified random sampling method was used to extract the portion from the training data, which was selected from classification data[13]. Eventually, the total 3500 samples were randomly selected. The total amounts of urban areas from these classified satellite images can be used as the global constraint for urban simulation.

4.2 Data-mining of transition rules and urban simulation

By the principle of ACO, one classification rule corresponds to each ant route, and discovering classification rules can be regarded as searching optimal routes by ants. So the spatial variables are treated as the attribute nodes of ant route, and the cells, whether they have been translated into land use of urban, are treated as class nodes of ant route. If a cell has been converted to urban development, it can be marked as 1, while a cell that has not been converted to urban development can be marked as 0. Prior to rule discovery, continuous spatial variables must be discretized. Each route corresponds to one classification rule, and the discovery of a classification rule can be regarded as searching of optimal route by ants. The transition rules for urban development will be automatically derived from GIS and RS data through Visual Basic 6.0 programming With the training data selected above, 78 rules were yielded, part of which are listed in Table 2.

Table 1 Spatial variables required for derivation of transition rules using Ant-Miner

Spatial distance variables						Local variables	
Distance to city proper (PropD)	Distance to town centres (TownD)	Distance to national highways (NatD)	Distance to roads (RoadD)	Distance to railways (RailD)	Distance to expressways (ExprD)	Number of developed cells in the neighborhood (Nsum)	Constraint condition

Table 2 Part of the transition rules derived by using Ant-Miner

Rule 1:
IF
RoadD<=23 and ExprD<=25 and 181<TownD<=206 and land use='cropland' and Nsum >= 3
Then
Converted to urban development (confidence =0.92)
Rule 2:
IF
NatD<=47 and RoadD<=23 and 23<ExprD<=102 and 153<RailD<=179 and TownD<255 and land use= 'orchard' and Nsum >=4
Then
Converted to urban development (confidence =0.86)
Rule 3:
IF
413<PropD<=450 and 115<RoadD<=138 and RailD >230 TownD>235 and Nsum <2
Then
Not converted to urban development (confidence =0.83)

In order to validate the reliability of Ant-Miner model, the transition rules discovered with Ant-Miner will go through accuracy test (Table 3). It is found that the accuracy is 74.6% for land that has been converted to urban development, is 79.3% for land that has not been converted to urban development, and the total accuracy is as high as 77.2%, which has already satisfied the requirement as far as the complex geographical data are concerned. In order to further validate the Ant-Miner model, a comparative study was carried out by using the See5.0 decision tree model. This is because the form of rule discovered from the See5.0 model does not show significant difference from that discovered from the Ant-Miner model, thus a meaningful comparison can be made between these two models. A comparison about the experimental results (Table 3) demonstrates that the total accuracy for the Ant-Miner model is nearly 5 percentages higher than that for See5.0 model, which indicates that the Ant-Miner model is more reliable and more suitable for knowledge discovery based on complex geographical data.

Table 3 Comparison of accuracy for Ant-Miner model and See5.0 decision tree model

Model	Accuracy of developed land (%)	Accuracy of undeveloped land (%)	Total accuracy (%)
Ant-Miner	74.6	79.3	77.2
See5.0	68.7	76.1	72.3

Based on the transition rules obtained from Ant-Miner, Simulation on spatial evolution of Guangzhou City during the period of 1988—1993 and 1993—2002 was ultimately implemented with ArcObjects plus Visual Basic 6.0. The whole process initially started from classification of TM data in 1988, the land use in 1993 and 2002 was then simulated by running this model with 200 iterations and 400 iterations respectively (Figure 4).

5 Model validation and comparison

Validation is usually required when the CA model is applied to the simulation of real cities[25]. A simple method to assess the goodness-of-fit is to compare the simulated patterns with the actual ones visually for validating CA[8,26], which is a rather preliminary method to validate the accuracy of the model. The visual comparison indicates that the simulated patterns are very similar with the actual patterns, which are obtained by the classification of remote sensing images (Figure 4).

Visual comparison is a rather preliminary method to validate the accuracy of the model. A further quantitative analysis is to produce a confusion matrix about the concordance between the simulation results and the actual urban patterns. It is based on the spatial overlay of the simulated and the actual development patterns cell by cell. Table 4 lists the comparison of these two patterns in 1993 and 2002 for the ACO-based CA model. The total accuracies are 83.3% and 76.8%, and the Kappa coefficients are 0.64 and 0.53 for the simulation of urban development in 1993 and 2002, respectively. The simulated results in 1993 have a better accuracy because it uses the case closer in time.

Structural conformity is also important in the assessment of simulation results[12]. The indicator of Moran I can be used as the spatial statistics for measuring the patterns. Moran I is a useful spatial indicator that can reveal the degree of spatial autocorrelation[27]. The indicator is able to estimate how close the simulated land use pattern is to the actual urban development[12]. Table 5 shows the structural conformity using the indicator of Moran I. The Moran I values are 0.627 and 0.687 for the simulation of land development in 1993 and 2002 respectively by using ACO-CA model. They are 0.626 and 0.684 for the actual land development in 1993 and 2002 respectively. This indicates that there is a good conformity between the simulated and actual land development according to the measurement using the indicator of Moran I.

In order to further validation the model, transition rules derived from the See5.0 decision tree model are used to simulate the urban development in the study area,

Table 4 Simulation accuracies of different model for the Guangzhou

Model	Year	Total accuracy (%)	Kappa coefficient
ACO	1988—1993	83.3	0.64
	1988—2002	76.8	0.53
See5.0	1988—1993	81.5	0.60
	1988—2002	73.2	0.46
Null model	1988—1993	82.2	0.59
	1988—2002	67.8	0.35

Table 5 Assessment of the goodness-of-fit for the CA model using Moran I index

Time	1988	1993	2002
Actual	0.633	0.626	0.684
Simulated (ACO)	0.633	0.627	0.687
Simulated (See5.0)	0.633	0.621	0.680

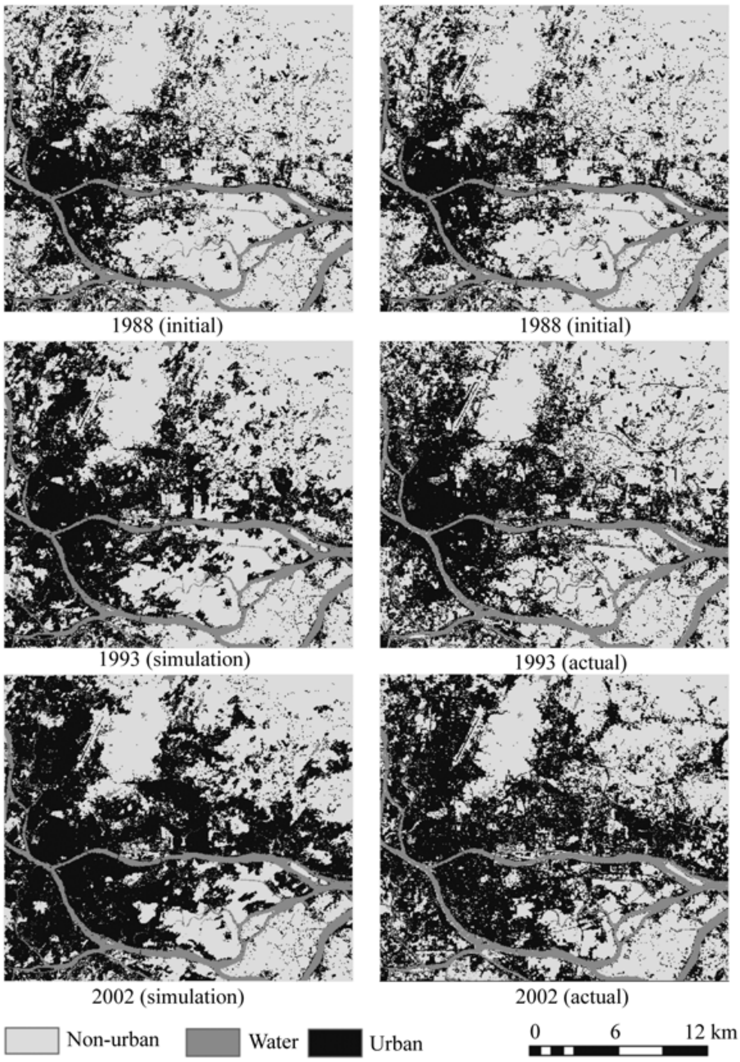

Figure 4 The simulated and actual urban development in Guangzhou in 1988, 1993, and 2002.

and the accuracy of simulation results based on cell by cell comparison is calculated and listed in Table 4, which shows that the total accuracies of simulation are 81.5% and 73.2%, and the Kappa coefficients are 0.60 and 0.46 for the simulation of urban development in 1993 and 2002, respectively. Table 4 indicates that the ACO-based CA model has much better simulation performance than the See5.0-based CA model.

Recently Pontius and Malanson pointed out that a predictive model should be compared with a Null model of pure persistence (no change) for model validation[28]. A Null model is a kind of model that predicts nothing as nothing would change. The baseline is that a predictive model should have better performances than a null model[28]. For instance, urban land use changes 15% in a period, so a Null model of pure persistence would be 85% correct based on the standard overall accuracy, while the overall accuracy of a predictive model should be higher than 85%. However, the overall accuracy has a bias because of the difference between the actual agreement and chance agreement[29], which can be effectively explained with Kappa coefficient, especially under the condition of geographical system with position, number and integrated information. As a result, more meaningful results will be yielded through comparing the Kappa coefficient between the Null model and a predictive model[25], Kappa coefficient is calculated as

$$\text{Kappa} = \frac{M\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+}\cdot x_{+i})}{M^2 - \sum_{i=1}^{r}(x_{i+}\cdot x_{+i})}, \tag{11}$$

where x_{ii} are the elements on the main diagonal of the

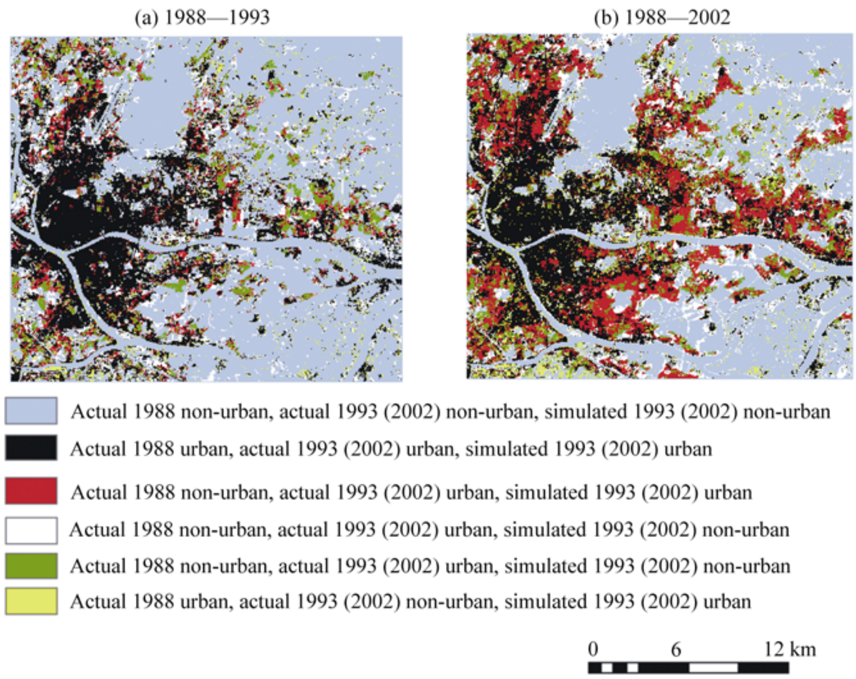

Figure 5 Distribution of agreement and disagreement of the simulated patterns or urban development of Guangzhou in 1993 and 2002.

error matrix, x_{i+} is the sum of the ith row of the error matrix, x_{i+} is the sum of ith column of the error matrix. The Kappa coefficient of the Null model is listed in Table 4. In the period of 1983—1993, the total accuracy of ACO-CA model is only 1.1% higher than that of Null model, but the Kappa coefficient is 0.05 higher than that of Null model, demonstrating a larger difference between Kappa coefficient; In the period of 1988—2002, the total accuracy of ACO-CA model is 9% higher than that of Null model, while the Kappa coefficient is 0.18 higher than that of Null model, demonstrating a remarkable difference in accuracy, which indicates that the ACO-CA model is rather a powerful tool for simulating urban development.

Figure 5 further displays the spatial distribution of agreement and disagreement of the simulated patterns of urban development in Guangzhou in 1993 and 2002. Correct simulation results from ACO-CA model are displayed with grayish blue, black and red color, while the incorrect are shown as other color parts, and correct simulation results from Null model are also displayed with the grayish blue, black and blue parts.

6 Conclusions

Simulation of complex resources and environment system is not only theoretically significant but also shows high prospect in application. Complex system has posed some problems for models based on traditional equations, which cannot well meet with the simulation of resources and environment system, while the CA model based on ‘bottom-up’ approach proves effective on simulating the evolutionary processes of complex systems, and is also widely applied to geographical phenomena, many significant research results have since been achieved. The key to CA model is how to define transition rules, however, most of which are defined by mathematical equations, and are hence not explicit, and it is difficult to determine the parameters involved in these formulae.

In this paper the ACO-based geographical CA (ACO-CA) was proposed. ACO is actually a complex multi-agent system, composed of a great number of artificial ants with simple intelligence. Complex tasks, such as optimized route for seeking foods, can be effectively fulfilled by the mutual cooperation between ants. As this type of ‘bottom-up’ approach coincides well with CA, which is, therefore, very appropriate for discovering CA transition rules. In our work data mining technique deriving classification rules from ant colony optimization (Ant-Miner) was first introduced into geographical CA. ACO was used to construct CA transition rules, which are not expressed in mathematical formulae and more easily comprehended by people, so can describe the complex relationships in a more convenient and precise way. With strongly robust, self-adaptive, nonlinear and characterized by positive feedback mechanism, ACO is more powerful to extract rules for complex geographical phenomena.

The ACO-CA model is applied to Guangzhou City, remote sensing images of different years were used as the major observation data. Stratified random sampling method was used to extract the portion from the training data for discovering transition rules, and Ant-Miner was used to automatically mine CA transition rules. In this way the urban development of Guangzhou in the period of 1993—2002 was simulated. Compared with the real situations of Guangzhou cell by cell, the accuracy and the Kappa coefficient of ACO-CA model are high.

Further comparison between ACO and See5.0 decision tree model also indicate that ACO-CA model shows higher accuracy in simulating urban development. This could be due to the fact that the pheromone in Ant-Miner is updated continuously and the positive feedback it provides can help to correct any mistakes resulting from the defects of heuristic function. The entropy involved in decision tree method is a measure of partial heuristic and may easily be affected the interactions among attributes. Particularly, in dealing with data of strong correlation among attributes, the decision tree method could be easily trapped in the problem of partial optimization of space searching. However, the pheromone in Ant-Miner is based on the overall performance of a rule, so the dynamic updating of pheromone can better deal with data correlation.

1 Couclelis H. Cellular Worlds: a Framework for Modeling Micro-Macro Dynamics. Environ Planning A, 1985, 17: 585—596[DOI]
2 Couclelis H. Of mice and men: what rodent population scan teach us about complex spatial dynamics. Environ Planning A, 1988, 20: 99—109[DOI]
3 Couclelis H. Macrostructure and Microbehavior in a MetropolitanArea. Environ Planning B, 1989, 16: 141—154[DOI]
4 Clarke K C, Brass J A, Riggan P J. A cellular automata model of wildfire propagation and extinction. Photogramm Eng Rem S, 1994, 60: 1355—1367
5 White R, Engelen G. Cellular automata and fractal urban form: a cellular modelling approach to the evolution of urban land-use patterns. Environ Planning A, 1993, 25: 1175—1199[DOI]
6 Batty M, Xie Y. From cells to cities. Environ Planning B, 1994, 21: 531—548
7 Zhou C, Shun Z, Xie Y. Geographic Cellular Automata Study (in Chinese). Beijing: Science Press, 1999
8 Clarke K C, Hoppen S, Gaydos L. A self-modifying cellular automaton model of historical urbanization in the San Francisco Bay area. Environt Planning B, 1997, 24: 247—261[DOI]
9 Li X, Yeh A G O. Modelling sustainable urban development by the integration of constrained cellular automata and GIS. Int J Geogr Inf Sci, 2000, 14(2): 131—152[DOI]
10 Wu F. SimLand: a prototype to simulate land conversion through the integrated GIS and CA with AHP-derived transition rules. Int J Geogr Inf Sci, 1998, 12(1): 63—82[DOI]
11 Li X, Yeh A G O. Cellar automata for simulating complex land use systems using neural networks. Geogra Res (in Chinese), 2005, 24(1): 19—27
12 Wu F. Calibration of stochastic cellular automata: the application to rural-urban land conversions. Int J Geogr Inf Sci, 2002, 16(8): 795—818[DOI]
13 Li X, Yeh A G O. Neural-network-based cellular automata for simulating multiple land use changes using GIS. Int J Geogr Inf Sci, 2002, 16(4): 323—343[DOI]
14 Li X, Yeh A G O. Knowledge discovery for geographical cellular automata, Sci China Ser D-Earth Sci, 2005, 48(10): 1758—1757
15 Liu X P, Li X. Retrieving CA nonlinear transition rule from high-dimensional feature space. Acta Geogr Sin (in Chinese), 2006, 61(6): 663—672
16 Colorni A, Dorigo M, Maniezzo V, et al. Distributed optimization by ant colonies. Proceedings of the 1st European Conference on Artificial Life, 1991, 134—142
17 Dorigo M. Optimization, learning and natural algorithms. Ph. D. Thesis, Department of Electronics, Politecnico diMilano, Italy, 1992
18 Dorigo M, Maniezzo V, Colorni A. Ant system: optimization by a colony of cooperating agents. IEEE Trans Systems, Man, and Cybernetics-Part B, 1996, 26(1): 29—41[DOI]
19 Kwang M S, Weng H S. Multiple ant-colony optimization for network routing. Proceedings of the First Int Symposium on Cyber Worlds, 2002, 277—281
20 Lumber E, Faieta B. Diversity and adaptation in populations of clustering ants. Proceedings of the Third International Conference on Simulation of Adaptive Behavior: From Animals to animates, MIT Press/Bradford Books, Cambridge, MA, 1994, 501—508
21 Duan H B. Ant Colony Algorithms: Theory and Applications (in Chinese). Beijing: Science Press, 2005
22 Parpinell R S, Lopes H S, Freitas A A. Data mining with an ant colony optimization algorithm. IEEE T Evol Comput, 2002, 6(4): 321—332 [DOI]
23 Brewlow L A, Aha D W. Simplifying decision trees: a survey. The Knowl Eng Rev, 1997, 12(1): 1—40[DOI]
24 Wu F, Webster C J. Simulation of land development through the integration of cellular automata and multicriteria evaluation. Environ Planning B, 1998, 25: 103—126[DOI]
25 Li X, Liu X P. An extended cellular automaton using case-based reasoning for simulating urban. Int J Geogr Inf Sci, 2006, 20(10): 1109—1136[DOI]
26 Ward D P, Murray A T, Phinn S R. A stochastically constrained cellular model of urban growth. Computers, Environ Urban Systems, 2000, 24: 539—558[DOI]
27 Goodchild, M. Spatial Autocorrelation, CATMOG, 47, GeoBooks, Norwich, 1986
28 Pontius G. R, Malanson J. Comparison of the structure and accuracy of two land change models. Int J Geogr Inf Sci, 2005, 19(2): 243—265 [DOI]
29 Congalton R G. A review of assessing the accuracy of classification of remotely sensed data. Remote Sens Environ, 1991, 37: 35—46[DOI]