

## The Research and Analysis of Association Rules on the Internet of Things

Xie Yanxin

Jilin Agricultural Science and Technology College  
Jilin 132001, Jilin, China  
e-mail: xieyanxin\_8@163.com

Project supported by the Jinong Institute word[2012]

422

Wang Yongjiang

Jilin Agricultural Science and Technology College  
Jilin 132001, Jilin, China  
e-mail: 187121686@qq.com

**Abstract**— Temperature, humidity, and light intensity are the important influence indicators of crop production. In order to achieve good economic and social benefits, we could use of modern network technology (3G network, wireless sensor network, mobile network, and local area network), sensing technology, RFID technology, database technology and mobile computing technology to build Internet platform. We can monitor crop growth process in real-time, increasing the industrialization level and crop yields, improving the ability to resist natural disasters, extending the crop growth time, etc. Therefore, with the rapid development of the Internet, China's agriculture development would have a new stage.

**Keywords**- the Internet of Things; network technology; real-time monitoring

### I. THE INTERNET OF THINGS

In *Future Path*, Bill Gates refers to the concept of the Internet of Things as Internet connecting individual things. Through the combination of radio frequency identification (RFID) device and intelligent computing technology (sensing equipments for information such as infrared sensors, global positioning system, laser scanner etc.), the Internet of Things, in accordance with certain contract, makes objects of the real world connect each other, with the help of Internet, exchanging correspondent information and communication, and achieving all-round interconnect between equipments.

### II. THE APPLICATION OF NETWORKING TECHNOLOGIES AT HOME AND ABROAD IN AGRICULTURE

On August 7, 2009, premier Wen Jiabao examined of research and development center of sensor network engineering technology one branch of the Chinese Academy of Sciences, in Wu Xi, He spoke highly of the innovation of opening up "perception China" (also called "wisdom China").

Nowadays, when the agricultural production in China transit to the intelligent, by applying the modern network technology (3G network, wireless sensor network, mobile

network, and local area network), sensing technology, RFID technology, database technology and mobile computing technology to build Internet platform. Therefore, with the rapid development of the things Internet, China's agriculture development would have a new platform[1].

China's agricultural development mainly falls into the three areas of the field (crops), fruit trees and vegetables. In Jilin province, there mainly exists three kinds of crop production, ie. corn, rice and soybean production. Moreover, greenhouse production has been lagged by various factors such as geographical position of Jilin and the native climate.

In the 1950 s, tunnels technology (simple facilities agriculture) was introduced to China from the Soviet Union. Today, China has started extensively introducing foreign large even and supporting the cultivation techniques in greenhouse, China's agriculture began to adopt large-scale, intensive and scientific production and management methods.

The Internet of Things technologies were first applied in the cultivation of Rome greenhouse (from BC three years to AD 69, the application of mica covering for generating precocious cucumber [2] was recorded). The Netherlands, Germany, Japan, South Korea, and the United States etc. are leaders in the world on the overall level [3].

### III. INTELLIGENT GREENHOUSE STRUCTURE OF THE INTERNET OF THINGS

The intelligent greenhouse control of The Internet of Things is a huge automatic control network system, which first uses sensors in a space distribution wireless network[4] of autonomy, of which each node equipped with sensors perceiving environment factors of plant growth, controlling-node device equipped with execution to control execution devices to improve environment factors of plant growth. The sensor-node and controlling-node cooperate with each other and provide appropriate environment for plants' growth. Secondly the system uses close RFID communication technology, receiving the parameters of the environmental factors of objects.

The figure of the wireless sensor network system demonstrates intelligent greenhouse control of the Internet of Things ("Figure 1"). The wireless sensor network has of five nodes, which are four sensor-nodes and one controlling node. The first node is temperature and humidity collection equipped with temperature and humidity sensor node, real-time monitoring the internal temperature of the greenhouse. The second node is light acquisition nodes using photoconductive resistance to detect the situation of internal light in the greenhouse. The third node is when someone appears around the greenhouse. security-node sends information to the main controlling center, and a light alarming. The fourth node is monitoring the thick of carbon dioxide in the air, since photosynthesis of plant needs light and carbon dioxide. When the level of light reaches the system setting, the system will automatically open the fan to strengthen ventilating, providing enough carbon dioxide for plants. The last node is humidity acquisition node (normally used together with the first node). The infinite sensor networks system is nerve ending of the Internet of Things, responsible for detecting parameter changes of the entire environmental factors, and sends the data collected through the coordinator to the other end of the network of processing machine of users.

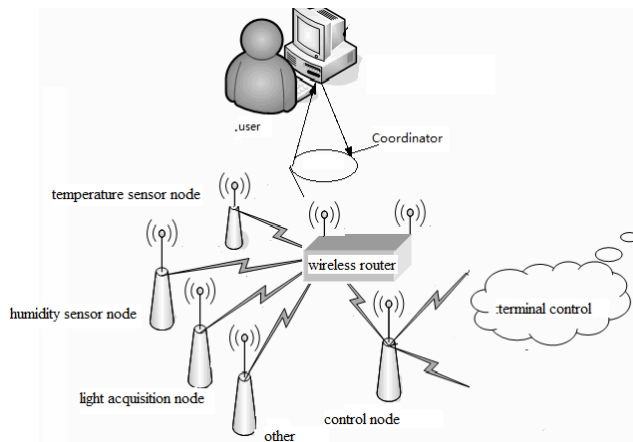


Figure 1. the wireless sensor network system in the intelligent greenhouse control system of The Internet of Things

#### IV. PROBLEMS

The greenhouse development, consists many problems, such as the influence of environmental factors, low level of facilities, poor ability of environmental control, low degree of mechanization, lagged standards and norms and lagged theory and technology research. And the gap to intelligent agriculture and machinery agriculture is large. The development of greenhouse is strictly restricted by factors such as human labor and costs.

Now we use computer technology, sensing technology, network technology, and other methodologies to regulate internal environmental factors (such as temperature, humidity, light, the CO2 etc.) [5] in greenhouse. We also rely computers more dynamically to control environment factors in greenhouse for crop production, however, in practice, the collected data from each node is sent to users, and according to the information users sent back, we control other control nodes.

Although we reach the purpose of controlling, we do not make comprehensive evaluation of existing data and collected data, nor make mass data analysis, so we do not make full use of each node of the collected data comprehensive control, then this therefore data mining technology also depend on the historical data and new acquisition mass data analysis.

#### V. ANALYSIS ON PARAMETERS OF ENVIRONMENT FACTORS BY ASSOCIATION RULES

The Internet of Things is applied to fine agriculture. Since the Internet of Things combines with the Internet and mobile networks, it can comprehensively "percept" the parameter index of environmental factors, at any time and anywhere and then it shifts from "perception" stage to "control" stage. How to analyze the parameters index of the sensor? We try to use association rules to analyze. The mining of association rules is one of the important issues of the data mining. The association rules use in greenhouse control, The purposes are digging out the hidden relationship between data of those levels of temperature, humidity, light and CO2. From the group of given set of data items, it is mining interesting, frequently occurred connection mode between parameters of environmental factors [6]. Through the mining of association rule, we can find a certain inner links between items of database to effectively control the automatically controlling and deciding ability of indoor crops. Apriori algorithm is a classic association rule mining algorithm, which is groundbreaking to use the technology based on support for pruning, systematically controlling the growth of candidate item sets index.

We can use RFID, sensors and two dimensional barcode to get environmental parameters at anytime and anywhere. In the following table 1-1, G1, G2, G3, G4 and G5, are on behalf of the orders of environment factors such as the level of temperature, humidity, light, CO2, H2O. Each record shows the group of data of environmental factors, namely an affairs record, which consists of data items of environmental factors, as seen in table 1. According to the relationship between index parameters, we can find out the relevant control measures the control node should take.

TABLE I. PARAMETERS OF ENVIRONMENT FACTORS

NO.	environment factors
001	G1, G2, G5
002	G2, G4
003	G2, G3
004	G1, G2, G4
005	G1, G3
006	G2, G3
007	G1, G3
008	G1, G2, G3, G5
009	G1, G2, G3

1) First calculate C1, which is the 1- item set, scanning all transaction record, and statistic the data items that meet the demand of support request.

2) C1 carries on the self-connection operation, and the statistic the support, then we can get the result as following,

$$C2 = \{ (G1, G2), (G1, G3), (G1, G5), (G2, G3), (G2, G4), (G2, G5), (G3, G5) \}$$

3) From the self-connection operation by C2, we get  $L3 = \{ \{ G1, G2, G3 \}, \{ G1, G2, G5 \}, \{ G1, G2, G4 \}, \{ G2, G3, G4 \}, \{ G2, G3, G5 \}, \{ G2, G4, G5 \} \}$ . By now, we can use the Apriori property to get the judgement that the latter 4 candidate items is impossible to be frequent, and we can may remove directly, therefore we get  $C3 = \{ \{ G1, G2, G3 \}, \{ G1, G2, G5 \} \}$

4) Computing  $L4 = \{ \{ G1, G2, G3, G5 \} \}$  further. Because the subset  $\{ I2, I3, I5 \}$  is not frequent, the algorithm runs to the end, C4 is empty set and the frequent set is C3.

Transform parameters of environment factors into boolean matrix ("Figure2"):

$$D = (D_1 \ D_2 \ D_3) = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Figure 2. boolean matrix

## VI. ASSOCIATION RULES PRODUCED BY FREQUENT ITEM SETS

For each frequent item sets L, calculate the reliability respectively on every true subset according to the following formula, then further determine the condition items and the items been determined(1):

$$\frac{\text{support\_count}(l)}{\text{support\_count}(s)} \geq \text{min\_conf} \quad \square \square (1) \square$$

If it can meet the above conditions, output the rule "S  $\Rightarrow$  (l-s)". In this rule, min\_conf is the minimum confidence threshold. Take the frequent item sets {G1, G2 and G5} as an example, its nonvoid subsets are {G1, G2}, {G1, G5}, {G2 and G5}, {G1}, {G2} and {G5}

The results association rules are as follows, each has been listed its confidence.

$$G1 \wedge G2 \Rightarrow G5, \text{confidence} = 2/4 = 50\%$$

$$G1 \wedge G5 \Rightarrow G2, \text{confidence} = 2/2 = 100\%$$

$$G2 \wedge G5 \Rightarrow G1, \text{confidence} = 2/2 = 100\%$$

$$G1 \wedge G2 \Rightarrow G5, \text{confidence} = 2/6 = 33\%$$

$$G2 \Rightarrow G1 \wedge G5, \text{confidence} = 2/7 = 29\%$$

$$G5 \Rightarrow G1 \wedge G2, \text{confidence} = 2/2 = 100\%$$

Through the comparison with minimum confidence parameters, we can determine the results get from association rules.

From this example we can see, association rules get through the calculation of confidence may be more, but as to the parameters of greenhouse, not all associations have meanings, the challenge is to judge the reasonable ones and the deceptive ones from numerous association on the basis of their experience and expertise. So how to select the valuable association, and how to provide the interesting (valuable) association rules are critical questions in the association analysis system. Therefore, some have introduced the concept of interestingness, which defined from the objective perspective called the objective interestingness, while the ones defined from user's side are called subjective interestingness. The measures mainly used in this paper is to select association rules by using the objective interestingness, whose definition and measure are data driven. We define the formula of interestingness as follows:  $RI = P(X, Y) - P(X)P(Y)$

In this formula, P(X) is the probability of X's appearance, similarly, P(Y) is the probability of Y's appearance, and P(X, Y) is the probability X and Y the appears at the same time. When "RI=0", the results show that X and Y are independent of each other, when  $RI > 0$ , the results indicates that the antecedent and the seccedent have a positive correlation, that is, the emergence of X will increase the probability of Y appearance. When " $RI < 0$ ", the results indicates that the antecedent and the seccedent have a negative correlation, that is, the emergence of X will decrease the probability of Y appearance. The system selects the association rules that have positive correlation as output.

## VII. FUTURE PROSPECTIVE OF THE INTELLIGENT GREENHOUSE

1) It can make planting with high-tech developing high-tech products of vegetables, or planting some rare flowers and trees. Currently, overall development of technology greenhouse in domestic is far lower than abroad. So high-tech intelligent greenhouse in China is just beginning but its speed of development is quite rapid. The main reason is the domestic requires the technology of the intelligent greenhouse.

2) The ecological restaurant has gradually risen in domestic and is accordance with the modern concept of life: green, environmental protection, healthy.

3) The combination of intelligent greenhouse and flower market have already formed in the scale in foreign countries and also started to build in China markets, which have splendid future for the next 10 years.

4) Move part of the field crop breeding to the greenhouse.

5) The control of greenhouse environment and the management of crop cultivation are developing into

intelligentized and networked, Greenhouse industry is transferring to the area that are energy saving and low cost, so the energy saving technology is becoming the focus of research.

## REFERENCES

- [1] Song Junde, et al; 《Wireless communication and networking》; Modern data communications and wireless communications series; Beijing — National Defense Industrial Publishing House, 2010.02
- [2] International Telecommunication Union UIT. ITU Internet Reports 2005: The Internet of Things[R], 2005, 112.
- [3] Gan Zhixiang. The origin and development of Internet of Things[J]. Modern Economic Information, 2010, (1) : 158, 157.
- [4] Dae-Man Han, Jae-Hyun Lim. Design and implementation of smart home energy management systems based on ZigBee. IEEE Transactions on Consumer Electronics . 2010
- [5] Wang Janzhou. 《The “Internet of things” will become a driver of economic development》 [J]. IT Time Weekly, 2009(10):20.
- [6] Harald Sundmaeker, Patrick Guillemin, et al. Vision and Challenges for Realising the Internet of Things. 2010