1125

A New Resource Information Integrating Method in Semantic Concept Networks

Li Sheng Qi

College of Economics & Management, Zhe Jiang Agriculture and Forest University Ling-an City, China, postal codes: 311300 e-mail: sqli65@163.com

Abstract

Information integrating in Semantic Concept Networks (SCN) is one of the important operations in these networks. These operations require resource usage. Due to the limited resource of units, the resource productivity should be considered as a key objective in design of Semantic Concept Networks. Therefore the Integrating is a suitable method that used in resource usage management. For this purpose many methods have been proposed. In this paper, the idea of not sending time correlated information of units has been considered by using the sequential function. Also, a model to estimate the remaining resource of units by the base station has been presented. The proposed ideas have been implemented over the LANDING-C protocol. Evaluation results showed that the proposed methods had a better performance in resource usage and the lifetime of the network in comparison with similar methods.

Keywords Integrating, Semantic Concept Network, Information Interaction, AAD

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Semantic Concept Networks are a class of wireless ad hoc networks. In these networks, Semantic Concept units *collect* information from physical environment and after processing sent to the base station (BS1). Thus allow monitoring and control many types of physical parameters. Each Semantic Concept unit has limited resource and in most applications, replacing resource sources are not possible. So lifetime of Semantic Concept units is highly dependent on resource stored in their battery. Integrating is a designing method that used for management of Semantic Concept Networks. In this method, the network is divided into several independent collections that these collections called cluster. So each cluster contains a number of Semantic Concept units and a cluster head unit. Member units in a cluster send their information to relative cluster head unit. Cluster head unit aggregates these information and send to the base station. Therefore, Integrating in Semantic Concept Networks has advantages such as information aggregation support [1], information integrating facilitation [2], organizing a suitable structure for scalable routing [3], and utilization propagation of data in the network [4].

Information integrating in Semantic Concept Networks is an important operation in these networks and for this purpose many methods have been proposed. The LANDING2 protocol has been considered as a hierarchical basic method. This method is suitable for monitoring applications. Each unit periodically senses the information and sends them. In this method, the Integrating method has used for information integrating and aggregation. The cluster and cluster head selected randomly, therefore there is no assurance to select the exact improved number and uniform distribution of cluster head throughout the network. Many improvements in LANDING protocol have been presented. LANDING-C3 method is an example of these improvements. In *LANDING*-C, the forming of clusters is done using a centralized method by the base station in the starting of each period. Base Station uses the received information from units for finding the predetermined number of cluster heads and network configuration within the clusters. This information contains position and resource of units. Another improvement to this method is the use of estimation. One of these methods is LANDING-CE4. In the proposed technique resource level collected from all units in two primary periods but not collected in the other periods. Instead, the average resource of initial periods is

used. Considering that the resource estimation in this method is not precise, this Integrating scheme is not utilization and suitable. There is some proposed Integrating methods that ECVB are samples of them.

Each Semantic Concept unit is observer of a physical phenomenon. Also physical phenomenon such as temperature and continuously change in time. So the information provided by Semantic Concept units is dependent on each other. Some methods that based on not sending of correlated data are considered. The VINA5 method is one of them. In this method the Semantic Concept unit compares the value of sampled data with previous data, if that be different send it and otherwise goes to sleep mode. The proposed improvement to this method is that Semantic Concept unit decides to send data with comparing the value of new sample with last reported data. These mentioned methods due to error in report, is not suitable. Therefore, a method proposed to increase the accuracy of data reporting. For precise estimation of unit resource, the base station must be aware of data time Interaction. So with existence of data time Interaction and using resource estimation of units, a method suggested so that the base station can estimate units resource precisely. These methods avoid the overhead excess and increase the network lifetime.

2. Related Work

2.1 LANDING

One of the most famous hierarchical routing protocols based on Integrating, is the LANDING protocol. In this method, each cluster members send their data to cluster head. The cluster head aggregate this data and send to the BS. So the communication cost is reduced. The operation of cluster forming and data transmission in LANDING is done in two phases. Setup phase is the stage of forming cluster and cluster head. At this stage, cluster and cluster head randomly selected. After forming the cluster, cluster head propagate TDMA6 scheduler to specify the data transfer time to member units. Then the steady-state phase started. In the steady-state phase, each member unit in cluster send data to the cluster head only in its time slot and at the rest of time pieces for resource conservation goes to sleep mode. In this method, the cluster head between network units. Many improvements over the LANDING method have been provided that in these improvements firstly, as far as possible the best Integrating and cluster head selection is done, secondly possible as possible overhead of the protocol is to be reduced. LANDING-C method is an example of these improvements.

2.2 LANDING-C

In LANDING-C, clusters forming in the beginning of every period are done, using the centralized method by the base station. The base station uses received information from units that includes resource and unit status, uses this information during the setup phase for finding predetermined number of cluster heads and network configuration within the clusters. Next classification of units in the clusters is done to minimize resource usage in order to transfer their data to the related cluster head. Results show that LANDING-C overall performance is better than LANDING because of the optimal forming of clusters by the base station. In addition, the number of cluster heads in each period of LANDING-C is equal to the expected optimal value. While in LANDING the number of cluster heads varies in different periods because of lack of global coordination. As in LANDING-C at the beginning of every period resource of units must be sent to BS, therefore units early discharged and the network lifetime reduces. Another improvement on this method is the use of resource estimation. The LANDING-CE method is an example of these methods.

2.3 LANDING-CE

In the LANDING-CE method, the resource level of all units collected only in two primary periods and not be collected in other periods. Instead because of knowing information about resource level of units, we can calculate resource usage average for each unit by using information of two primary periods. This means that reducing calculated resource level from the resource level of unit, causes predicting current resource level of unit. The problem of this method is that firstly resource estimation is not done precisely and secondly if units have

correlated data, while not sending correlated data means that previous data is valid, so this plan of Integrating is not suitable and utilization.

2.4 VINA

Phenomenon that observed by Semantic Concept units, continually change in time. Therefore information received by the units is correlated on each other. These cases for physical phenomena that are continuous or repetitive, or in an application that the accuracy is not too important, or in a network that unit density in a region is high, have seen more. There are two types of data Interaction: 1) spatial Interaction; 2) Time Interaction. In the spatial Interaction, aggregation is done within the network by cluster heads. This is one of the proposed methods to reduce resource usage. So the units that have correlated data send them to cluster head and cluster head after aggregating these data send to the base station. This causes to prevent waste of resource. This method has been implemented in LANDING protocol.But in the time Interaction; each unit can have correlated data in successive times. Wang Wei proposed the VINA method. The main idea of VINA method was that the Semantic Concept units send their data only when this data differ with previous data otherwise goes to sleep mode. This method has a reporting error. There is an improvement to this method that presented below.

2.5 Improvement of VINA

In this method, the Semantic Concept unit decides to send data by comparing the value of newly sampled data with last reported data. However, Semantic Concept units maintain last reported data. For better understanding, we describe this section with an example. Suppose that a given Semantic Concept unit that received data are 1.0, 0.95, 1.05, 0.95, 1.05 respectively. A threshold has been considered that data changing to this threshold is not important. The value of this threshold considered equal to 10%. First given data that is equal to 1.0 successfully sent and in the next period 0.95, will not be sent when:

$$\frac{|0.95 - 1.0|}{1.0} = 5\% < 10\%$$

5% < 10% Otherwise that will be sent.

In the third stage $\frac{|1.05 - 1|}{1} = 5\% < 10\%$ that will not sent and in the fourth stage

 $\frac{|0.95 - 1.0|}{1.0} = 5\% < 10\%$ that will not sent. This method is suitable when phenomenon changes have of phenomenon change continuously with time. So most of data are in ascending or descending mode in the time slices. Or in an application such as temperature for example in a certain time slot occurs a specific event. So the proposed methods have errors and are not suitable. We offer a method to improve this method and prevent the waste of resource. In addition, the problem of data time Interaction is not considered in proposed protocols. Therefore, we will check the time Interaction of data in the proposed methods.

3. The Proposed Method

3.1 Presentation of Methods

Three ideas are proposed here: 1) the data time Interaction; 2) Resource estimation model of units and 3) the hybrid method.

In the data time Interaction method, Sequential function Analyzing method (TSA) used to decide sending or not sending of data. Then in time t in the beginning of each period, base station send percentage of error e(t) to all units. First data sensed by unit and sent. Second and third and fourth data sent based on the improved VINA method. Then the unit runs sequential function function to determine the value of forecast of trend line model, to create trend model. In the next times the sensed data compared with predicted value of trend model, if the difference between these two values exceeds a threshold value, data sent to the given unit and the recalculate forecast function of trend model to update the trend line. Otherwise, the Semantic Concept unit does not send the sensed data with this insurance that sensed data placed in accuracy range of data. So only some data have to be sent that are very different from the trend line model. This help to prevent resource loss.

We call the unit resource forecast model LANDING-CEC8 and describe as follow. For doing the best integrating, that is needed to know resource of the units. The estimation method is a method that has low cost and is suitable. We also use the resource estimation method. For this, we divided LANDING-C protocol to three phases. To pology building phase, setup state and steady-state. In the first phase units send their position to the BS. Then BS creates network topology based on these positions. Once the topology was formed in the base station, base station unit calculates the distance of units to each other. The BS calculates the amount of resource used in each unit in the setup-phase, using a simple mathematical model. Then deduct this amount from primary resource and calculates its remaining resource. Finally do the Integrating and goes to the steady-state phase. In this phase for each unit, the data time correlated method applied according to the following method.

BS unit should be informed of data time Interaction in units to estimate precisely resource of them. Therefore cluster head create a c that containing list of all members of the cluster. Cluster head registers every unit in to the table that have correlated data and do not sent in certain times. In the end of each period, cluster head sends this table with collected data to the base station. This table contains units ID and number of times that these units not sent data. Base station uses this information for Integrating decisions in centralized methods. Ultimately that cause to resource estimation in centralized methods is more carefully while the best Integrating is created and the network lifetime increases. So in total lifetime of the network, first phase has done once but setup and steady-state phases done as in LANDING-C.

3.2 Process of Proposed Methods

Linear forecast method is a powerful technique to predict sequential function in an environment changing with time. Suppose that you want to contact an independent variable x and a dependent variable y to specify. If we assume that the true relationship between these variables in a straight line and the value observed for each value of y for every given x is a random variable then we can wrote:

$$E(y \mid x) = a_o + a_1 x \tag{1}$$

where in this equation a_0 is the width from the origin and a_1 is the slope of the line that is unidentified fixed values. Observed value y can be described with the following equation where the error ε created because of not conforming real value to the amount of predicted value.

$$y = a_0 + a_1 x + \varepsilon \tag{2}$$

This pattern is usually named a simple linear regression model. Because that has only one independent variable so that the x independent variables called forecast variable and y called the response variable. Forecast and response variables x and y can be sequential function in which case we have a sequential function regression pattern. There are several methods to estimate unknown parameters a_0, a_1 in Equation (2) that can be used. One a method that a lot used is the least square error method in which the a_0, a_1 estimates obtained from minimizing asume of squares errors or remaining's. Suppose we have n observations of $(x_1, y_1), (x_2, y_2) \cdots (x_n, y_n)$. A model that is based on these observations is written as follows:

$$y_i = a_0 + a_1 x_i + \varepsilon_i, i = 1, 2, \cdots, n$$
(3)

And the total square error is as the follow:

$$\phi(a_0, a_1) = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$$
(4)

So the total square error is simply the total squares of deviations observed y_i and $a_0 + a_1 x_i$. The estimated values of a_0 and a_1 that we call them a_0 and a_1 that achieved using the least squares method by minimizing $\phi(a_0, a_1)$ toward a_0 and a_1 so we can write:

$$\frac{\partial \phi}{\partial a_0} = -2\sum_{i=1}^n (y_i - a_0 - a_1 x_i) = 0 \qquad \qquad \frac{\partial \phi}{\partial a_1} = -2\sum_{i=1}^n x_i (y_i - a_0 - a_1 x_i) = 0$$

This system of equations called least squares line normal equations system that simplified as the following:

$$na_0 + a_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$
 $a_0 \sum_{i=1}^n x_i + a_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i$

By solving this system a_0 and a_1 estimates or on the other hand a_0 and a_1 obtained and so:

$$\hat{a}_{1} = \frac{\sum_{i=1}^{n} x_{i} y_{i} - \left[\sum_{i=1}^{n} x_{i}\right] \left[\sum_{i=1}^{n} y_{i}\right] / n}{\sum_{i=1}^{n} x_{i}^{2} - \left[\sum_{i=1}^{n} x_{i}\right]^{2} / n} = \frac{\sum_{i=1}^{n} (x_{i} - \overline{x})(y_{i} - \overline{y})}{\sum_{i=1}^{n} (x_{i} - \overline{x})^{2}} = \frac{s_{xy}}{s_{xx}}$$

$$\overline{s}_{i} = \frac{1}{2} \sum_{i=1}^{n} x_{i} - \frac{1}{2} \sum_{i=1}^{n} y_{i}$$

where

e $x = -\sum_{i=1}^{n} x_i$ and $y = -\sum_{i=1}^{n} y_i$, So the fitted simple linear regression model is as follows: $\overline{y} = \hat{a}_0 + \hat{a}_1 x$ For each value of the predicted variable x we can obtain corresponding

value predicted response from this equation. The fitted values of \hat{y}_i corresponding to observed values \hat{x}_i for every i = 1, 2, ..., n is the following: $\overline{y}_i = \hat{a}_0 + \hat{a}_1 x_i$

The difference between ith given fitted value to an observed y_i value called a residue so that:

$$\hat{e}_i = y_i - \hat{y}_i, i = 1, 2, \cdots, n$$
 (5)

If the fitted model regression for data is appropriate, in this case remains do not follow of appropriate form. There is no clear pattern for the remains. This method stated by Equation (11) that is a recursive method:

Linear forecast method is a powerful technique for predicting sequential function in a time-varying environment. This method is expressed in Equation (6) and is a recursive method:

$$y(t+T) = a_1 y(t) + a_2 y(t-T) + \dots + a_m y(t-(m-1)T)$$
(6)

Estimated value at time t as a linear function of previous values in the times "t - T, t - 2T, \cdots , t -

mT" has been produced is obtained. In Equation $(1) a_1, a_2, \dots, a_m$ are the linear forecast coutilizations, "m" is the model degree, "T" is the sampling time, y(t + T) is the next observation estimation and y(t), y(t - T), ..., y(t - mT) are the present and past observations. The forecast error which is the difference between the predicted and the real values (Equation (7)) must be minimized.

$$Error(\%) = \left\{ \frac{pridicted \quad value - \text{Re}al \quad value}{\text{Re}al \quad value} \right\} * 100\%$$
(7)

In order to estimate the coutilizations of linear forecast model we use the least squares error method and rewrite Equation (6) with considering modeling error in Equation (8):

$$y(t) = a_1 y(t - T) + a_2 y(t - 2T) + \dots + a_m y(t - mT) + e(t)$$
(8)

The error e(t) is generated because of not adopting the linear forecast model to the real value. So to find the coutilizations, a_1, a_2, \dots, a_m in Equation (8) we use the sum least squares error and set of linear functions. Presented in Equation (9).

$$\begin{bmatrix} y(t)\\ y(t-T)\\ \vdots\\ y(t-mT) \end{bmatrix} = \begin{bmatrix} y(t)y(t-2T)\cdots y(t-mT)\\ y(t-2T)\cdots y(t-(m+1)T)\\ \vdots\\ Y(t-(k+1)T)\cdots y(t-(m+k)T) \end{bmatrix} * \begin{bmatrix} a_1\\ a_2\\ \vdots\\ a_m \end{bmatrix} + \begin{bmatrix} e(t)\\ e(t-T)\\ \vdots\\ e(t-kT) \end{bmatrix}$$

$$Y = \omega \times A + E$$
(10)

Elements in the matrix A are the coutilizations which can be found by least squares error method (11):

$$A = (\varphi^T \varphi)^{-1} \varphi^T Y \tag{11}$$

In Equation (11), φ^{T} , is the transpose of the matrix φ , and $(\varphi^{T}\varphi)$ is the inverse of matrix. In practice:

If m is chosen larger than is required (i.e. overestimation of the model order), Equation (11), cannot be solved for any unique set of coutilizations, because of some columns in the matrix ϕ ,

are not independent of each other. Hence $(\varphi^T \varphi)$ would be unique and will not have inverse. This means that the system of equations in Equation (8) will have an infinite number of answers for the coutilizations. Geometrically speaking, it is like fitting an infinite number of lines to a single point which is not the preferred case.

If m is chosen less than the required value, the number of independent equations would be more than the number of unknown variables $(a_1 - a_m)$. Such a system of equations has to be solved for the best approximation of coutilizations. The best approximation for coutilizations $(a_1 - a_m)$ is the use of the least squares error method. Obviously, that is how much precision, the number of the data sent will increase and vice versa.

As previously mentioned, LANDING-C protocol divided to three phases: topology building phase, setup-state and steady-state and this method is proposed as following. These three phases is called LANDING-CEC.

a) Topology building phase

1) Start of network.

2) Base station receives position of all units that contain x and y.

3) Base station calculates distance of all units with each other's using the following Formula (12):

$$d_{ab} = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$$
(12)

4) End of phase.

b) Setup-state phase

5) If Integrating has changed

5.1) BS calculate resource usage per unit using the Formulae (13) and (14):

$$E_{Tx}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{friss-amp}d^2 : d < d_{crossover} \\ lE_{elec} + l\varepsilon_{two-ray-amp}d^4 : d \ge d_{crossove} \end{cases}$$
(13)

$$E_{Rx}(l) = lE_{elec} \tag{14}$$

And after computing, place the sum of two variables in to the E_{ni}^{\prime} .

5.2) Base station calculates remaining resource (E_R) of each unit using the Formula (15).

$$E_{R} = E_{R} - E_{ni}^{\prime} \tag{15}$$

5.3) BS audit the information received from CHs, if the units name with the number of duplicate data is there, considered in terms of computing units remaining resource.

5.4) If $E_R \leq 0$ the unit is dead.

5.5) Otherwise the unit is alive and participates in Integrating.

6) Integrating is formed.

 E_{ni}^{t} : The amount of consumed resource by unit i in time t.

c) Steady-state phase(Hybrid method)

7) If unit n is a cluster head.

7.1) Cluster head creates a table containing unit name and the number of correlated data.

7.2) Cluster head collects data from the members of the cluster.

7.3) If the cluster head has not received data from its member unit, registers name of that unit in the table and a unit to be added to the number of correlated data. Cluster head sends the table with correlated data to the base station.

8) Otherwise, that is a member unit.

8.1) If the unit is in turn on then the unit runs correlated data method.

8.2) Otherwise goes to sleep mode.

 $E_{_{TX-amp}}(l,d)$: Strengthening the resource to transferring 1 bit data in distance d.

 $\mathcal{E}_{friss-amp}$: Radio resource of amplifier. d: distance between receiver and sender, l: length of data package.

Table 1 should be created by the cluster head in the steady-state phase.

Table 1. Information of node's correlated data in BS.

Node name	Number of data correlated
Node2	2
Node5	1

4. Simulation and Evaluation of Methods

4.1 Simulation Environment

Simulations have been done on the Redhat9.0 Linux operating system by using NT2 network simulator. LANDING and LANDING-C protocols Implementation are from the CAMP project at SJTU (Shanghai Jiaotong University) on NT2.

Simulations have been done on the dimensions of 100×100 m area with 100 Semantic Concept units. The primary resource of every Unit is 2 joule. During our simulations we have chosen some random network topologies to obtain the mean results. Two modes for base station location have been considered. One position in (50 and 50) the exact center of the network and the latter is in position (50 and 100), which is near the area under monitoring. Each period of simulation takes 20 seconds long. Receivers and transmitters follow the model that their parameters are:

$$E_{elec} = 5.0 \times 10^8 \quad J/bit, \quad E_{tx} = E_{rx} = 5.0 \times 10^8 \quad J/bit$$

$$\varepsilon_{free-space-amp} = 1.0 \times 10^{-11} \quad J/bit/m^2$$

$$\varepsilon_{two-ray-amp} = 1.3 \times 10^{-15} \quad J/bit/m^4$$

$$E_{two} = E_{tx}$$

 L_{tx} , L_{rx} are Send and receive power needed for each bit. Simulations have done using LANDING, LANDING-C, LANDING-CE and LANDING-CEC protocols.

Simulation assumptions:

1) All units are static and have limited resources.

A New Resource Information Integrating Method in Semantic Concept Networks (Li Sheng Qi)

2) Base station has not limited resources.

3) All units at any moment have data to send.

4) All Units equipped with the location determination

4.2 The Result of Simulation

In the NT2 simulator and also LANDING and LANDING-C protocols, data produced with the Uniform distribution. But in fact phenomenon that seen by Semantic Concept units are continuously changing with time. Therefore, the information received by the Semantic Concept units is correlated. Therefore generated data by the simulator must have a normal distribution. Normal distribution Definition: we say a random variable x is normally distributed if its density is as follows:

$$f_x(x) = f_x(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2} / 2\sigma^2$$
(16)

where the parameters μ and σ are $^{-\infty < \mu < +\infty}$ and $\sigma > 0$. Each distribution with the given density functions as defined in relation (16), called a normal distribution. To show the parameters we have used μ and σ^2 symbols because we have known that these are the mean and variance parameters of distribution respectively. For producing data by normal distribution we assume $\mu = 0.8$ and $\sigma = 1$. In the mentioned protocols data generated in uniform distribution. So we have changed the code of these protocols to normal distribution. In addition, the actual amounts of resource in each unit in all protocols of LANDING, LANDING-C, LANDING-CE and LANDING-CEC in each period was calculated. In our first scenario where the base station along the network are located at the point (50 and 100).

Before reviewing mentioned protocols we first survey the data Interaction protocols. We will examine VINA, improvement of VINA Protocols and the idea proposed in conjunction with data Interaction (TSA). The produced Data by a unit in the methods VINA, improvement of VINA and the proposed method (TSA) is reviewed. We took this result that the number of data submissions in TSA method is less than previous methods and the accuracy is high. We have run each of desired protocols 20 times, so that resulting graphs are the average of results of the runs. Then we calculated the mean of results and by running data Interaction methods on them we extracted the following results. We concluded that the number of sent data in TSA method is less than two other methods and have a high accuracy. Error Percentage for correlated data in TSA method is equal to 1%. In a period in specified times we have generated 20 times random data and we repeat it again 20 times. Then we calculated the average of them and have run the data Interaction methods over them and extracted the following results.

5. Conclusion

This article solves the problem of correlated data in all of discussed protocols in this paper. So the units that have time correlated data and sending this data wastes their resource and thus network lifetime will decrease. By using the method of data time Interaction, the problem will be raised. Also we have eliminated periodic sending of units data in LANDING-C protocol. By using resource estimation in LANDING-CEC method there is no need for units to send their resource level and position to the base station. They only have to send their position at the beginning of network to the base station and the base station creates network topology and using a simple mathematical calculation will calculate the resource of units.

References

- M Yokoo and K Hirayama. "Channel Access for Cellular Mobile Systems Using Constraint Satisfaction Techniques". *Principles and Practice of Constraint Programming*. Lecture Notes in Computer Science. 1999; 1713: 490-491.
- [2] A Gamst. "Some Lower Bounds for a Class of Channel Access Problems". IEEE Transactions on Vehicular Technology. 1986; 35(1): 8-14.
- [3] JK Hao, R Dorne and P Galinier. "Tabu Search for Channel Access in Mobile Radio Networks". *Journal of Heuristics*. 1998; 4(1): 47-62.

- [4] KN Sivarajan, RJ McEliece and JW Ketchum. "Channel Assignment in Cellular Radio". Proceedings of 39th IEEE Vehicular Technology Society Conference, San Francisco. 1989; 5(1): 846-850.
- [5] J Cardoso, A Sheth, J Miller, J Arnold and K Kochut. "Quality of Service for Workflows and Web Service Processes". Web Semantics: Science, Services and Agents on the World Wide Web. 2004; 41(3): 281-308.
- [6] Ruo Hu. "Stability Analysis of Semantic Concept Network Service via Data Stream Methods". Applied Mathematics & Information Sciences. December, 2012.
- [7] Ruo Hu. " Channel Access Controlling in Semantic Concept Network using Smart Grid System". *Applied Mathematics & Information Sciences*. December, 2012.