A Kind of Context-aware Computing Approach for Proactive Service

ZHANG De-Gan[1, 2] HUANG Xiao-Bin[3]

Abstract We focus on modeling and computing of aware context with uncertainty for making dynamic decision during seamless mobile service. We re-examine formalism of random set, which is not finite-set statistics (FISST), argue the limitations of the direct numerical approaches, give new modeling mode based on random sets theory (RST) for aware context, and propose our computing approach of modeled aware context. In addition, we extend classic D-S evidence theory after considering reliability, time efficiency, relativity of context, and compare these two kinds of relative computing methods for uncertain context. By comparing, the validity of new context-aware computing approach based on improved random set theory (IRST) or extended D-S evidence theory (EDS) for proactive service has been tested.

Key words Proactive service, context-aware, random set theory, D-S evidence theory

1 Introduction

In order to realize seamless mobility of pervasive computing, context-aware process must be considered for attentive seamless service, because context-aware information during seamless transfer is helpful for reasoning, making decision, and realizing service in time[1]. Random sets theory (RST) is one theory of applied mathematics, which is not finite-set statistics (FISST)[2]. We re-examine formalism of random set, argue the limitation of the direct numerical approaches, and propose our computing approach. D-S evidence theory is another method for expressing and computing context. In order to ensure the QoS of proactive service, we will modify the computing method of evidence after considering context′s reliability, time efficiency, and relativity. We call the modified method extended D-S evidence theory (EDS), which has improved the classic fusion rule of D-S evidence theory.

2 Modeling based on IRST

Using the notion of the Janossy density[3], we can define the joint probability density of two random finite sets, X and Y, and the conditional probability density such as $P(X|Y)$ and $P(Y|X)$. Suppose that X is a finite random set modeling the unknown number of objects to be estimated and Y is an observation with respect to X given as another finite random set. If the Jonossy density is jointly defined for the two random sets, X and Y, then we can apply Bayes′ rule as

$$P(X|Y) = P(Y|X)P(X)/P(Y) \tag{1}$$

which gives us the formal "answer" to the multi-object estimation problem that is defined by the object model $P(X)$ and the observation model $P(Y|X)$. Assuming that, for the moment, "objects" are all static, a typical model X for objects is a Poisson point process with an intensity measure G on the state space E. In order to define a multi-sensor, multi-U scan problem, let us consider N observations that are given as finite random sets, $(Y_1, Y_2, \cdots, Y_N)$, in measurement spaces, $(E_1, E_2, \cdots, E_N)$, each having σ finite measure μ_k. We assume conditional independence of observations as

$$P((Y_k)_{k=1}^N|X) = \prod_{i=1}^N P((Y_K)|X) \tag{2}$$

Then the problem can be defined completely when we specify each measurement model $P((Y_K)|X)$. A typical model, assuming: 1) object-wise independent detection, 2) object-wise measurement mechanism, and 3) independent Poisson point process modeling false alarms, can be written as

$$P((Y_k)|X) = \mathrm{e}^{-\nu_k} \sum_{a(X,Y_k)} (\prod_{x(a)} p_m(a(x)|x)p_D(x))\times$$
$$(\prod(1-p_D(x)))(\prod(r_k(y))) \tag{3}$$

As a conditional Janossy density, where $p_m(y|x)$ is the density of the object state to measurement transition probability, $p_D(x)$ is the probability of an object at state $x \in E$ being detected (included) in the observation Y_k, R_k is the density of the intensity measure of the Poisson point process modeling false alarms in Y_k with $\nu_k = \int_{E_k} r_k(y)\mu_k(\mathrm{d}y)$, and $A(X, Y_k)$ is the set of all the one-to-one functions defined on a subset Dom(a) of X taking values in Y_k, then, for any integer $k' \le N$, there is a collection $E_{k'}$ of a collection λ, called data to data association hypotheses, of tracks, each of which is a subset of the tagged cumulative data sets, $U_{k=1}^{k'} Y_k \times \{k\}$.

3 Computing of modeled aware context

According to the modeling based on random set theory for aware context, the solutions to the state estimation problems of a general class of multi-object were developed using the random finite sequences, *i.e.*, random point process formalism. Bayes′ equation (1) is solved numerically. The state space I of the collection of finite sets is approximated or truncated as $\bigcup_{n=0}^{n'} E^n$ or equivalently

$$I_{n'} = \{X \in I|\ \omega(x) \le n'\} \tag{4}$$

with a priori bound n' on the number of objects. The space E for each object state must be quantized in some efficient way.

Without any approximation (truncation), the cardinality of the collection I of all the finite sets in E, *i.e.*, the system state space, can be expressed as (By $n = \omega(A)$, we mean that the cardinality of set A is n.)

$$\omega(I) = \sum_{n=0}^{\infty}(\omega(E)^n)/n! = \exp(\omega(E)) \tag{5}$$

When repeated elements are not ignored but the orders in sequences are ignored, we consider quotient spaces of the direct product space E^n induced by permutations of elements. When the object state space E is finite, we have $I = 2^E$, which is the power set of E, and hence, $\omega(I) = 2^{\omega(E)}$.

Received January 4, 2006; in revised form August 22, 2006
Supported by National Natural Science Foundation of China (60503024, 60375038, 60374032) and Open Fund of Key Laboratory of Industrial Controlling Technology, Zhejiang University (060004)
1. Department of Computer Science and Technology, Tianjin University of Technology, Tianjin 300191, P. R. China 2. Key Laboratory of Industrial Controlling Technology, Zhejiang University, Hangzhou 310058, P. R. China 3. Beijing Bestpower Electrical and Engineering Ltd., Beijing 100096, P. R. China
DOI: 10.1360/aas-007-0860

In a sense, the direct numerical calculation approach in fact replaces the curse of hypotheses explosion in (4) by the curse of dimensional explosion. A solution to this problem of dimensional explosion was proposed in [4], in which, for each (hypothesized) number $n (\leq n')$ of objects, the density functions $p'_{n_1}, p'_{n_2}, \cdots, p'_{n_n}$ of n a posteriori probability distributions on the object state space E, together with the posteriori probability on the number of objects, $q'_1, q'_2, \cdots, q'_n$, approximate the posteriori Jonassy density function as

$$P\left(X|(Y_k)_{k=1}^{k'}\right) = q'_{\omega(x)} \cdot \sum_{a \in A(x)} \prod_{x} p'_{\omega(x) a(x)}(x) \tag{6}$$

where $A(X)$ is the set of all the one-to-one functions defined on set X in E taking values in the set $\{1, \cdots, \omega(X)\}$ of integers. For each (hypothesized) number n of objects, the joint probability of the states of n objects is approximated by n independent probability distributions in (6). In other words, for each n, all the associated hypotheses are "combined" and cross-correlation among n objects are ignored, *i.e.*, for a given n, a non-Gaussian extension of the algorithm, known as joint probabilistic data association (JPDA) algorithm is used. With this approximation, computational complexity can be bounded by $(n'(n'+1)/2)\omega(E)$.

We can modify (4) to accommodate merged measurement as

$$P\{Y|\{x_1, x_2, \cdots, x_n\}\} = p_C(x_1, x_2) \cdot \sum_{y} p_{CM}(Y|(x_1, x_2)) \cdot p_{CD}(x_1, x_2) p_{FA}(Y \backslash \{y\}) + (1 - p_C(x_1, x_2)) \cdot \sum_{a(\{1,2\}, Y)} \cdot \left(\prod_{i(a)} p_m(y|x) p_D(x)\right) \cdot \left(\prod (1 - p_D(x))\right) \cdot p_{FA}(Y(a)) \tag{7}$$

where $p_C(x_1, x_2)$ and $p_{CD}(x_1, x_2)$ are the probability of two objects at x_1 and x_2 being merged and that of the merged measurement being included in the data set Y, respectively, $p_{CM}(\cdot|\cdot)$ is the density of joint-where-state-to-measurement transition probability, and $p_{FA}(\cdot)$ is the Janossy density of the random set of false alarms. The object-wise detection probability p_D and the density of the state-to-measurement transition probability p_m are the same as [3].

Another interesting variation of the sensor model (3) may be a predetection tracking model[5], such as

$$P(Y|X) = P((y(j)_{j \in J})|X) = \prod_{j \in J} 1/(\sqrt{2\pi}\sigma(j)) \cdot \exp\left(-\frac{\left[\dfrac{(y(j) - \sum_{x \in X} S(j|x)}{\sigma(j)}\right]^2}{2}\right) \tag{8}$$

which is a conditional probability density of an observation $Y = (y(j))_{j \in J}$ as a collection of intensity values integrated within each quantized two or three dimensional cells, conditioned by the collection of objects modeled by a random finite set X. In (8), $S(j|x) = s(x) \int_j O(\eta - h(x)) \mathrm{d}\eta$ is the integrated contribution of an object at x within a cell, where $s(x)$ is the signal strength part of the object state x, $h(x)$ is the projection of the object state onto a focal plane or measurement space, $\Phi(\cdot)$ is an appropriate point-spread function, and $\sigma(j)$ is the standard deviation of the integrated noise in cell J, given cell-wise independent noises.

In what follows, we will discuss a potential new approach using random set formalism but without resorting to direct numerical calculation. The basic concept is borrowed from the stochastic clustering that Saha introduced in [6]. For example, in a single data set Y of a random finite set, the Choquet′s capacity functional of the random finite set X of objects can be written as

$$T(K|Y) = \{X \cap K \neq \emptyset \mid Y\} = 1 - \mathrm{e}^{-\int_E (1 - p_D(x))\mu(\mathrm{d}x)} \cdot \prod_{y \in Y} \frac{p_{FA}(y) + \int_{E \backslash K} p_m(y|x) p_D(x) G(\mathrm{d}x)}{p_{FA}(y) + \int_E p_m(y|x) p_D(x) G(\mathrm{d}x)} \tag{9}$$

For each K, where p_D and p_m are detection probability and transition probability density as described previously, and G is the intensity measure of the poisson point process X. (9) shows a cluster initialization process.

Now let us consider a cluster consisting of a set $\ddot{E}$ of hypotheses, each hypothesis $\ddot{e} \in \ddot{E}$ being a set of tracks τ, according to the notion of clustering in [6]. For each hypothesis λ, a posteriori probability $p'(\lambda)$ is attached, and for each track, a posteriori object state probability density $p'(x|\tau)$ on E is given. Then consider the expected number of objects in this cluster in the sense that

$$\nu' = \sum_{\lambda \in \Lambda} \omega(\lambda) p'(\lambda) \tag{10}$$

Let $T = \bigcup \lambda$ be the set of all the tracks in the cluster, and consider the "mean" probability density in the following sense.

$$p'(x) = \frac{\sum\limits_{\tau \in T} q'(\tau) p'(x|\tau)}{\sum\limits_{\tau \in T} q'(\tau)} \tag{11}$$

where $q'(\tau) = \sum \{p'(\lambda) | \tau \in \lambda \in \Lambda\}$ is the track probability for each track τ, defined using hypothesis probabilities $p'(\lambda)_{\ddot{e} \in \ddot{E}}$. For example, if the posteriori distribution represented by $p'(x|\tau)$ for each track has a finite sufficient statistics, *e.g.*, each $p'(x|\tau)$ is Gaussian, then we may be able to approximate (11) by a probability distribution with finite sufficient statistics, *e.g.*, mean vector and variance matrix.

4 Extended D-S evidence theory

Based on classic D-S evidence theory mentioned above, we will discuss EDS in the following[7].

Let the function mass $m(\cdot)$ be a certain evidence information (context-aware information). We can define the exchange form $\hat{E}$ of this evidence E, where Θ is defined as above, A_i is focus element $m(A_i) > 0$, and i is the number of focus number which satisfies the conditions

$$\hat{m}(A_i) = m(A_i), \quad A_i \neq \Theta \tag{12}$$

$$\hat{m}(\Theta) = m(\Theta) + (1 - \delta) \tag{13}$$

where $\delta \in [0, 1]$ is context reliability factor after assessment according to specified case, $\sum \hat{m}(A_i) \leq 1$, with $\hat{m}$ being the basic probability assignment. Then, E is called the original evidence, and $\hat{E}$ is mapped evidence of E.

Let the function mass $m(\cdot)$ be a certain evidence information (context-aware information) E at the time-point t_0. Then, we can define the exchange form of the function mass

$$\hat{m}(A_i,t)=\xi(t-t_0)m(A_i), \quad A_i \neq \Theta$$
$$\hat{m}(\Theta,t)=\xi(t-t_0)m(\Theta)+[1-\xi(t-t_0)] \qquad (14)$$

where $\xi(t-t_0)=\delta f(t-t_0)$, δ is reliability factor and $f(t-t_0)$ is function of time efficiency which is provided by the expert of the special field of the object of interest and can be tuned after being assessed. The form of this function of time efficiency is various and changeable. In different fields, the description may be different, such as subsection function and trigonometric function.

Let the energy function $\Psi(E)$ of an evidence E be defined as

$$\Psi(E)=\sum_{i=1}^{n(E)} m(A_i)/|A_i|, A_i \neq \phi \qquad (15)$$

where A_i is the set of focus elements, $|A_i|$ is the radix of A_i, $n(E)$ is the number of elements and their set of power, $m(A_i)=\hat{m}(A_i,t)/\xi(t-t_0), A_i \neq \Theta, m(\Theta)=\hat{m}(\Theta,t)-[1-\xi(t-t_0)]/\xi(t-t_0), \xi(t-t_0)\neq 0, \hat{m}(A_i,t), \hat{m}(\Theta,t)$, and $\xi(t-t_0)$ are as defined above.

If the function masses $m_1(\cdot)$ and $m_2(\cdot)$ are basic probability functions of two evidences E_1 and E_2, their focus elements are A_i and B_j, respectively. Obviously, some focus elements of E_1 and E_2 may be relative, and the relativity degree is decided partly by the number of focus elements and its basic probability assignment. So we define the relative degree as follows.

The coefficients of relativities μ_{12} (which ranges from E_1 to E_2) and μ_{21} (which ranges from E_2 to E_1) are defined, respectively, as

$$\mu_{12}=\varphi(E_1,E_2)\Psi(E_2)/(2\Psi(E_1))$$
$$\mu_{21}=\varphi(E_1,E_2)\Psi(E_1)/(2\Psi(E_2))$$

where $\varphi(E_1,E_2)$ is the relativity degree of evidences E_1 and E_2, which can be computed as

$$\varphi(E_1,E_2)=2\Psi(E_1,E_2)/(\Psi(E_2)+\Psi(E_1))$$

5 Computing based on EDS

Let the mass functions $m_1(\cdot)$ and $m_2(\cdot)$ be basic probability functions of two evidences E_1 and E_2 in the space U, $\{A_i\}$ and $\{B_j\}$ are sets of focus elements. Then, the context fusion computing method considering context relativity is as follows

$$\hat{m}(A)=\sum_{A_i\cap B_j=A} m'_1(A_i)m'_2(B_j), A\neq\phi,\Theta \qquad (16)$$

$$\hat{m}(\phi)=0, \quad \hat{m}(\Theta)=\sum_{A_i\cap B_j=\Theta}\left(m'_1(A_i)m'_2(B_j)\right)+\eta$$

where

$$m'_1(A_i)=\begin{cases} m_1(A_i)(1-\mu_{12}) & A_i\neq\Theta \\ 1-\sum_{A_i\subset\Theta} m_1(A_i) & A_i=\Theta \end{cases}$$

$$m'_2(B_j)=\begin{cases} m_2(B_j)(1-\mu_{21}) & B_j\neq\Theta \\ 1-\sum_{B_j\subset\Theta} m_2(B_j) & B_j=\Theta \end{cases}$$

$$\eta=\sum_{A_i\cap B_j=\phi} m_1(A_i)m_2(B_j)$$

In the following, we give the fusion computing method of n evidences under consideration of context reliability.

Similarly, suppose the function masses $m_1(\cdot)$, $m_2(\cdot)$, $\cdots$, $m_n(\cdot)$ are basic probability functions of n evidences in the space U, and that the mapped functions are $\hat{m}_1$, $\hat{m}_2$, $\cdots$, $\hat{m}_n$, respectively. Then, the context computing method $\hat{m}$ is

$$\hat{m}(A)=c^{-1}\sum_{\cap A_i=A}\prod_{1\leq i\leq n} m'_i(A_i), \quad A\neq\phi \qquad (17)$$

$$\hat{m}(\phi)=0, \quad \hat{m}(\Theta)=(\sum_{\cap A_i=\phi}\prod_{1\leq i\leq n} m'_i(A_i))+\eta$$

where

$$m'_i(A_i)=\begin{cases} m_i(A_i)(1-\mu_{i\times(n-i)}) & A_i\neq\Theta \\ 1-\sum_{A_i\subset\Theta} m_i(A_i) & A_i=\Theta \end{cases}$$

$$\eta=\sum_{\cap A_i=\phi}\prod_{1\leq i\leq n} m'_i(A_i), \quad c=\sum_{\cap A_i\neq\phi}\prod_{1\leq i\leq n} m'_i(A_i)$$

6 Tests and evaluation

As an experimental example of our active space[7], we want to identity a person based on computation of two kinds of context tracked by face recognition agent and voice recognition agent, and then track his/her activity. For 200 persons, the reliability factor of the voice recognition agent $\delta_1=0.8$ and the reliability factor of the face recognition agent $\delta_2=1$. According to the gathered voice of 200 persons, the decision made by the voice recognition agent is $m_1(S,Z)=0.875$, which is in accordance with the collected image information of 200 persons, and the decision by the face recognition agent is $m_2(S)=0.9$.

If we consider the approach ((16), (17)) based on EDS mentioned above, we can get the mass for decision of the person′s identity as follows:

$m_3(S)=\hat{m}_1\bigoplus\hat{m}_2(S)=0.63+0.27=0.9$. Probability of the person is S.

$m_3(S,Z)=\hat{m}_1\bigoplus\hat{m}_2(S,Z)=0.07$. Probability of the person is S or Z.

$m_3(\Theta)=\hat{m}_1\bigoplus\hat{m}_2(\Theta)=0.03$. Probability of the person is uncertain.

In the above $m_3(S)$ shows that the probability of the person′s identity is S.

If we consider the approach ((7), (11)) based on RST mentioned above, then we are generating aggregate statistics for a group of objects in context-aware computing. Because there is addition probability in $m_3(S,Z)$ and $m_3(\Theta)$, which means the additional information about S or Z, for both S and Z, the additional probability is 0.1. We can determine that the probability of S is 0.9, so the probability region of S is [0.9, 1]. The computation result is consistent with our experiences, and there is no conflict of ideas between the two, so the efficiency of context-aware computing approach based on RST can be relied upon.

In experimental examples, with the increase of random finite sequence, the mean error ratio will decrease, which ranges from 0.247% to 0.089%, but with evidence theory method it ranges from 0.298% to 0.125%. By comparison, the advantage of RST is apparent.

7 Conclusion

In order to solve the attentive service problem of pervasive computing paradigm, we have studied context-aware computing during seamless mobility based on RST and EDS. We have argued the limitation of the direct numerical approaches, proposed our computing approach of modeled aware context based on IRST, and extended D-S evidence theory after considering the context′s reliability, time efficiency, and relativity. The validity of our approach has been successfully tested using the experimental example.

References

1 Satyanarayanan M. Pervasive computing: vision and challenge. *IEEE Personal Communications*, 2001, **8**(8): 10∼17

2 Mori S. Multi-target tracking theory in random set formalism. In: Proceedings of 1st International Conference on Information Fusion. Las Vegas, USA, 1998. 209∼216

3 Mahler R. Random sets as a foundation for general data fusion. In: Proceedings of the Sixth Joint Service Data Fusion Symposium. Laurel, USA, 2002. 357∼394

4 Castro P, Munz R. Managing context data for smart spaces. *IEEE Personal Communications*, 2000, **7**(5): 44∼46

5 Donald J, Lin L. Inferring high-level behavior from low-level sensors. In: Proceedings of International Conference of Ubiquitous Computing. New York, USA, 2003. 132∼139

6 Saha R K, Chang K C. An efficient algorithm for multisensor track fusion. *IEEE Transactions on Aero-space Electronic Systems*, 1998, **34**(1): 200∼210

7 Zhang D G, Shi Y C, Xu G Y. Learning by seamless migration-kind of mobile working paradigm. In: Proceedings of the 3rd International Conference on Web-based Learning. Springer-Verlag, Beijing, 2004. 115∼122

ZHANG De-Gan Professor at Tianjin University of Technology. He received his Ph. D. degree from Northeastern University. His research interest covers perversive computing, artificial intelligence, and computer games. Corresponding author of this paper. E-mail: gandegande@126.com

HUANG Xiao-Bin Engineer at Beijing Bestpower Electrical and Engineering Ltd.. He received his Ph. D. degree from University of Science and Technology Beijing. His research interest covers robotics, artificial intelligence, and computer animation.