IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITB12607

This paper appears in the book, *Emerging Trends and Challenges in Information Technology Management, Volume 1 and Volume 2* edited by Mehdi Khosrow-Pour © 2006, Idea Group Inc.

Evidential Characterization of Uncertainty in Location-Based Prediction

Iftikhar U. Sikder, Department of Computer and Information Science, Cleveland State University, Cleveland, OH 44115, USA, T: (216)687-4758, F: (216)687-5448, i.sikder.csuohio.edu

ABSTRACT

This paper provides a novel approach to characterization of uncertainty measures in classification and prediction of complex spatial objects in data mining. The paper shows the semantic limit of uncertainty measure in classical probabilistic approaches and presents a formal approach to characterize uncertainty parameters from Rough set and Dempster-Shafer's evidence theory in spatial domain. We have developed a rough set and Dempster-Shafer's evidence theory based formalism to objectively represent uncertainty inherent in the process of service discovery, characterization, and classification. Rough set theory is ideally suited for dealing with limited resolution, vague and incomplete information, while Depster-Shafer's evidence theory provides a consistent approach to model an expert's belief and ignorance in the classification decision process. Integrating these two approaches provide a mathematically consistent and objective means to measure belief, plausibility, ignorance and other useful measures in spatial classification. Moreover, it provides predictive measures of uncertainty and thereby allows including the context of spatial neighborhood effects.

INTRODUCTION

In recent years, there has been increased interest in understanding uncertainty in various aspects of computing. Numerous formal and informal systems have been developed for measure of uncertainty to reason under conditions of ignorance. Traditional probability framework allows commitment of partial beliefs to hypothesis; however the remaining belief is assigned to its negation leaving no room for assigning beliefs to the alternative hypothesis, which is often counter-intuitive. The Bayesian approach to evidence combination ignores the quality of evidence and normalizes varied probability judgments in a specialized way. Not until after the Dempster's generalization of the Bayes' theorem (Dempster, 1967) and reformulation by Shafer (Shafer, 1976) the uncertainty research turned to a prolific direction, popularly known as Dempster-Shafer Theorem. In early 80's Pawlak introduced the rough set theory (Pawlak, 1982;) based on the notion of *indiscernibility* and *vagueness*. Later, it was found that the basic notions of evidence theory can be interpreted from the framework of rough set theory (Skowron and Grzymala-Busse, 1994).

On the other hand, many approaches of uncertainty have been proposed to characterize data and spatial processes (Mowrer and Congalton, 200). Implicitly, uncertainty semantics in spatial information is often associated with error, imprecision, reliability, validity, confusion, ignorance, noise and incompleteness. Although imprecision and imperfection is an endemic aspect of spatial information (Goodchild, 1995) traditional spatial models and computational algorithms rarely allow to incorporate them in the application modeling process.

EVIDENTIAL REASONING USING ROUGH SET MODEL

Unlike probability theory, evidence theory differentiates between ignorance and uncertainty. The classical probability theory does not allow the concept of imprecise probability measure. The axiom of additive law in probability theory requires sharply defined bound of probability. Dempster-Shafer theory replaces the classical concept of additive measure of probability by the concept of superadditives and subadditive. Conceptually, Dempster-Shafer theory rests on the lack of ability to map the objects and our knowledge of their attributes. In contrast, rough set is concerned with the mapping of decision concepts with respect to the attribute granules or equivalence classes. Different of perspectives of Dempster-Shafer's belief functions can be found in (Lingras & Wong, 1990). However, a formal integration of rough set and evidence theory was initially proposed by (Skowron & Grzymalla-Busse, 1994). They provided an interpretation of the belief and plausibility function in terms of the lower and upper approximation of rough sets. The interpretation was further extended by (Yao & Lingras, 1998).

SPATIAL NEIGHBORHOOD CONTEXT AND UNCERTAINTY MEASURE

In this research we try to map evidence theoretic measure of probability to sets of decision attributes of spatial objects to understand which decision attributes belong to which objects and at the same time estimate the contextual constraints of spatial classification of the neighborhood evidences of decision classes. This allows us to model mutually close neighbor in terms of *cliques*, which is defined as a subset of a neighborhood system in which all pairs of sites are mutual neighbor. Thus, the stochastic ordering of the spatial relation of the attributes of objects allows introducing the belief measure of the appropriateness of a class with respect to attribute and spatial context.

The Belief and Plausibility Functions for Decision Table

Let Θ be the frame of discernment such that Θ = {θ1,…,θk }and Ar = ($U_{\partial r}$, Ar ∪ {d_{∂}}) be a decision table for a given site. The belief function for the frame of discernment Θ and corresponding compatible decision table A can be expressed as follows:

$$Bel_A(\theta)\sum_{\Delta\subseteq\theta} m_A(\Delta)$$

The standard basic probability function is derived from the union of unique decision elements of the equivalence class constitute a function

$$m_A(\theta) = \frac{|\Phi_B(\chi(\theta)|}{|U|} \qquad (1)$$

where, $\Phi_B(\theta) = E!d[s]_B^d$, [s]Bd $= \bigcup d(s) : s \in [s]_B$.

The basic probability assignment mA(θ)is the ratio of the number of objects with the same frame of discernment corresponding to their equivalence class to the number of objects in the universe.

Let Θ be a frame of discernment, Θ = {θ1,…,θk }. The decision d in decision table determines another frame of discernment Θ_A where Θ_A = {1,…,r(d)}. Θ is compatible with A if |r(d)| = |Θ| or in other words the both frame of references have the same cardinality |Θ| = |Θ_A|. There exists a bijection $\chi(\theta)$: Θ → Θ_A between Θ and Θ_A such that $\chi(\theta)$ = i for i = 1,…,k. If s∈ [s]B then we can define the injection δA : U → P(Θ_A)

such that δA (s) is the unique subset θ of Θ_A where P(Θ_A) is the power set of Θ_A. For θ ⊆ the belief function is:

$$Bel_A(\theta) = \frac{\sum_{\Delta\subseteq\theta}|\Phi_B(\chi(\Delta))|}{|U|} \quad (2)$$

Let be the frame of discernment compatible with A = (U, A ∪{d}). For arbitrary θ ⊆ Θ the plausibility function is:

Following the definition of plausibility we have:

$$Pl_A(\theta) = 1 - Bel_A(\Theta-\theta) = 1 - \frac{\sum_{i\in\chi(\Theta-\theta)}|\Phi_B(\chi(\Delta))|}{|U|}$$

$$Pl_A(\theta) = \frac{\sum_{\Delta\cap\theta\neq\varnothing}|\Phi_B(\chi(\Delta))|}{|U|} \quad (3)$$

The Belief Functions and Plausibility Functions Generated by Neighborhood Decision Table

According to definition of the basic probability assignment for decision table

$$m_A(\theta) = \frac{|\Phi_B(\chi(\theta)|}{|U|}$$

For a given frame of discernment Θ = {θ1,...,θk }and the $\{d_\partial\}$) decision table in given site Ar = (, Ar ∪ {}). For simplicity, let us denote the $\Phi_B(\chi(\theta))$ as $\delta_\partial(\theta)$. For a given site r and clique c ∈ C (C1 ∪ C2 ∪ C3...) if we consider the non-empty intersections of the frame of discernments in a clique and normalize the sum of the cardinality of intersection so that $\sum_{\Delta\subseteq\Theta} m_r(\Delta) = 1$ and mr(∅) = 0 , then we get the following:

$$m_r(\theta) = \sum_{c\in C}\left|\bigcap_{\substack{\forall\partial_i\in c\\ \cap\theta_i=\theta}}\delta_{\partial_i}(\theta_i)\right| / |\Pi_\partial| \quad (4)$$

where, $|\Pi_\partial| = \sum_{c\in C}\left|\bigcap_{\substack{\forall\partial_i\in c\\ \cap\theta_i\neq\varnothing}}\delta_{\partial_i}(\theta_i)\right|$

For a given site r and clique c ∈ C (C1 ∪ C2 ∪ C3...) the belief function can be derived from the definition:

$$Bel(\theta) = \sum_{\Delta\subseteq\theta} m(\Delta)$$

Therefore,

$$Bel_r(\theta) = \sum_{\Delta\subseteq\theta}\sum_{c\in C}\left|\bigcap_{\substack{\partial_i\in c\\ \cap\Delta_i=\Delta}}\delta_{\partial_i}(\Delta_i)\right| / |\Pi_\partial| \quad (5)$$

The basic probability assignment is derived from the non-empty intersections of the frame of discernments in a clique. $|\Pi_\partial|$ is a normalizing function which ensure the property $\sum_{\Delta\subseteq\Theta} m_r(\Delta) = 1$ and mr(∅) = 0. For multi-site clique the basic probability assignment can be expanded as follows:

$$\{|\ \delta_r(\theta_1)\ |_{\substack{r\in c_1\\ \theta_1=\theta}} + |\ \delta_r(\theta_1)\cap\delta_{\partial 2}(\theta_2)\ |_{\substack{\{r,\partial_2\}\in C2\\ \theta_1\cap\theta_2=\theta}} +$$

mr(q)=

$$|\delta_r(\theta_1)\cap\delta_{\partial 2}(\theta_2)\cap\delta_{\partial 3}(\theta_3)|_{\substack{\{r,\partial_2,\partial_3\}\in C3\\ \theta_1\cap\theta_2\cap\theta_3=\theta}} + ...\}/|\Pi_\partial|$$

\+ (6)

Each term in the right hand side of the equation results from the intersections of the frame of discernment with respect to definition of clique.

In a similar manner we can show that $Pl_r(\theta) = \sum_{\Delta\cap\theta\neq\varnothing}\sum_{c\in C}\left|\bigcap_{\substack{\partial_i\in c\\ \cap\Delta_i=\Delta}}\delta_{\partial_i}(\Delta_i)\right| / |\Pi_\partial|$

(7)

Given the existing neighborhood evidence, we can derive the evidential measure of uncertainty using the above equations, which will help us to characterize a service with respect to certain predefined category. Thus, the decision regarding the appropriateness of services for a given purpose or allocating certain resources can be objectively measured with associated uncertainty hidden in the evidences. The same principle also applies to new services. Once we have defined the frame of discernment of the new services with respect to the indiscernibility relations of rough set, it is easy to extract the belief and plausibility functions from the neighborhood definition.

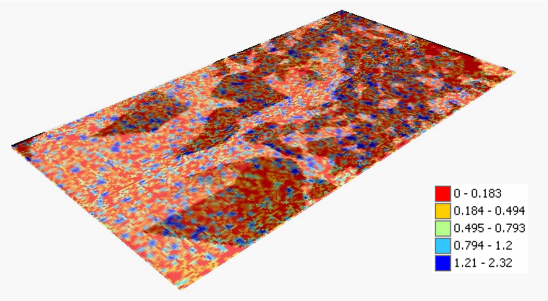
EMPIRICAL APPLICATIONS

We used a supervised classification in multiattribute geographic data. The decision classes are the land cover classes from satellite remote sensing images and the attributes used are biophysical evidences suggesting a land cover class. The experiment considers land cover classes as a focal element of a frame of discernment of all possible exhaustive classes (i.e., hypotheses) and the combination of biophysical attributes indicates the evidences in support of a set of land cover class for a given pixel. The evidential ambiguity of different classes is the result of spatial granularity of attributes, which is a typical characteristic of spatial processes. The concept of "evidence" in this case entails not only the attributes in support of decision classes, but also include spatial distribution of attribute sets and the context of neighborhood evidence with respect to given sample pixel.

Results: Randomness and Conflict in Evidential Claims

In a spatially homogenous region or in a spatially auto-correlated area, the degree of randomness in the distribution of spatial features was found relatively low. Inversely, in a heterogeneous area the candidate focal elements increase as a result of numerous equally likely decision categories. The overall distribution of Shannon entropy (Klir, 1994) in much of the region was found close to zero, because in many pixels there are only one focal element and the total mass assigned to the focal element is m ({x}) = 1. However, in most of these areas cardinality of the focal element is greater than one. Since, entropy measure is incapable of representing uncertainty resulting from nonspecificity of a focal element; apparently probability measure indicates that there is no uncertainty in these pixels. This is because probability measures are inherently fully specific and therefore incapable of characterizing the nonspecificity dimension of multisource information. Interestingly, the number of focal element is more than one in a thematic boundary regions and corresponding entropy score is higher in these areas. It was found that entropy measure performs well in characterizing boarder line regions. The nonspecificity score ranges from minimum value log2 |1| = 0 to the maximum possible log2 |5| = 2.3219 bits. One bit of nonspecificity displayed in the image express a total ignorance regarding the truth or falsity of the focal elements in the frame of discernment or possible subsets of decision classes. The result also shows a strong positive spatial autocorrelation (with a Moran's I of 0.94 and Geary's C value of 0.054 (Moran, 1950; Griffith, 1987). Since, nonspecificity

Figure 1. Spatial Distribution of Nonspecificity



of a pixel is estimated with respect to the neighboring pixel's frame of discernment within a context window, it is likely that nonspecificity score would reflect the affect of class boundary. In a boundary region, the total number of focal elements is greater than the inner region of a class. However, the cardinality of each focal element is not large enough. Therefore, in a boundary region the nonspecificity is low. Similar reasoning can be extended in the measures of discord. The discord function in the area shows a skewed distribution with a mean and standard deviation of 0.193 and 0.26 respectively. The total uncertainty is measured as the sum of discord and nonspecificity (Figure 1), which shows the same range as the nonspecificity and discord, i.e., $[0, \log_2|X|]$ or (0, 2.32190).

CONCLUSIONS

The uncertainty involved in the implicit stochastic effect of neighborhood evidences is formalized in a modified probability measure. Uncertainty measures established here incorporates estimation of randomness as well as conflicts in evidential claims of spatial as well as non-spatial evidences. The numerical uncertainty measures are primarily derived from the decision component of rough set equivalence classes, which are characterized by the attribute structure of the neighborhood context of spatial order. A key advantage of this model is that the model exploits the spatial coincidence or co-location association in the model induction process without introducing any subjective bias in the uncertainty measurement. The limitation of probabilistic measures are overcome by integrating rough set and Dempster-Shafer model in spatial prediction providing a means for improving classification schemes.

REFERENCES

Goodchild, M.F, and Buttenfield, B.P and Wood, J., (1995), I*ntroduction to Visualizing Data Validity*. In H.M. Hearnshaw and D.J. Unwin (eds) Visualization in GIS. New York: Wiley, 1995.

Dempster, A. P., (1967) *Upper and lower probabilities induced by a multi-valued mappin*g. Annals of Mathematical Statistics, 38:325-339.

Pawlak, Z. (1982). *Rough sets*. Internat. J. Comput. Inform. Sci., 11:341–356.Griffith, D.A. (1987), *Spatial Autocorrelation: A primer*. Association of American Geographers, Resource Publications in Geography.

Klir, G. J and Folger, T.A, (1988) *Fuzzy Sets , Uncertainty and Information*, Prentice-Hall, Inc., pp. 246-254,.

Lingras, P. J., & Wong, K. M. (1990). Two Perspectives of the Dempster-Shafer Theory of Belief Functions. *International Journal of Man-machine Studies, 33*, 467-487.

Moran, P.A.P. (1950), *Notes on continuous stochastic phenomena*. Biometrika, 37:17

Mowrer, H. T. and Congalton, R.G (2000). *Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing*. editors Ann Arbor Press, 2000.

Shafer, G., (1976) . *A Mathematical Theory of Evidence*. Princeton University Press, New Jersey.

Sikder, I., (2003), *Uncertainty Management in Empirical Knowledge Discovery and Classification in Spatial Databases, Doctoral Dissertation*. University of Maryland, Baltimore.

Smithson, M..(1989). *Ignorance and Uncertainty: Emerging Paradigms*. Springer Verlag, New York.

Skowron, A. and Grzymal³a-Busse, J. (1994). *From rough set theory to evidence theory*, in Advances in the Dempster-Shafer theory of evidence, John Wiley & Sons, Inc., New York.

Yao, Y. Y., & Lingras, P. J. (1998). Interpretations of belief functions in the theory of rough sets. *Information Sciences: an International Journal, 104*(1-2), 81-106.

Related Content