Visualizing the Uncertainty of Urban Ontology Terms

Hyowon Ban[1] and Ola Ahlqvist[1]

[1]Department of Geography, The Ohio State University,
1049B Derby Hall, 154 N Oval Mall
Columbus, OH 43210, USA.
{ban.11, ahlqvist.1}@osu.edu

Abstract. The concept of exurban is an example of a term likely to find its way into urban development ontologies. Many such terms are uncertain since there is no consensus of the exact definition of e.g. the boundary of exurbanization. In this research we focus on visualizing the spatial implications of the uncertainty in two existing definitions of exurban boundaries using empirical GIS data in Delaware County, Ohio, U.S. We argue that exurban boundaries are not crisp, hence, a series of fuzzy-set theory membership functions help define the uncertainty of the empirical exurban boundaries.

Keywords: uncertainty, visualization, exurban, boundary definition

1 Introduction

In Urban Civil Engineering (UCE) studies, there have been efforts to produce a taxonomy of ontologies, analyze the role of ontologies as a tool to foster an improved communication between stakeholders by building multi-lingual UCE glossaries of explication [24]. The studies emphasize that any serious attempt to construct urban ontologies must accommodate the evolution of concepts among different actors. This is because different groups have different concepts about the urban environment according to the inherently sociotechnical character of ontologies. Ontology also plays an essential role in the construction of Geographic Information Systems (GIS), since it allows the establishment of correspondences and interrelations among the different domains of spatial entities and relations [22].

Ontology, the science of being [21], is a logical theory accounting for the intended meaning of a formal vocabulary, and it determines what can be represented and what can be said about a given domain [24]. Fonseca et al (2000) analyze the urban environment from the ontologists' point of view. Other prospects suggested in the literature is composition of pre-existing independently developed ontologies, for instance, through the use of a context algebra to compose diverse ontologies [28], or through proxy contexts [2]. There are however a number of issues that make the application of ontology to urban areas problematic. Because of the differences in understanding concepts that form an ontology it is important that these differences can be articulated in some way. Some concepts in for example an urban area type ontology, terms such as urban, sub-/exurban, and rural areas, are inherently vague.

H. Ban and O. Ahlqvist: *Visualizing the Uncertainty of Urban Ontology Terms,* Studies in Computational Intelligence (SCI) **61**, 85–94 (2007)
www.springerlink.com

Yeates (1993) suggests 5 stages in the transition from exurban to suburban: agricultural, early urban influence, small town growth/exurbanization, and urban. Because the urban environment does not cease to exist abruptly—i.e., bona fide [20]—at the municipal borders, it is essentially fiat [20];[9] and should be treated as continuous.

Recently exurban areas have received specific attention because of its fast growth. According to Theobald (2005) in 2000, there were 125,729 km^2 of urban and suburban (<0.68 ha per unit) residential housing nationwide (conterminous USA), and about seven times that (917,090 km^2) of exurban housing (0.68–16.18 ha per unit). Statistics like these depend on definitions of what constitutes an exurban area and there are a number of suggested definitions. As there are various names to call exurbanization, most of them come with a separate definition. Irwin and Bockstael (2004) argue that since there is little consensus on a definition, data and measurements of sprawl are highly dependent on the researchers. Moreover, ontology is not likely to provide a 'silver bullet' simply because spatially continuous phenomena have received very little attention in the field of ontology [13]. There has been some research on boundary issues in exurban studies but few of them mention the uncertainty of the boundaries in exurban areas. Theobald (2005) models exurban land-use changes with a Landscape Sprawl (LS) metric. Wolman et al (2005) argues for measuring sprawl using data on density, concentration, centrality, nuclearity, and proximity of areas. Wilson et al. (2003) develop a model that determines the geographic extent, patterns, and classes of urban growth over time using land cover data. However, the exurban boundaries are defined crisp in their work. Caruso (2005) addresses the issue of urban expansion by exploring the emergence and morphology of a periurban zone at the periphery of a city, where residents and agricultural activities mix.

Generally uncertainty consists of errors, vagueness, and ambiguity [7]. Errors can be represented with probability, the vagueness can be explained by fuzzy set theory, and the ambiguity contains discord and non-specificity as its innate characteristics [7];[14]. Among the types of uncertainty, we think of vagueness when there is no unique allocation of individual objects to a class, or no precise spatial extent of the objects [7]. We can also find ambiguity when more than one definition for a term exist, one object is clearly defined but is shown to be a member of different classes under differing classification schemes or interpretations [14].

The purpose of this paper is to compare the spatial implications of different ontological commitments as they are represented by different definitions of exurban areas. We also want to demonstrate the relevance of representing exurban areas as vague objects by comparing the traditional crisp representation with a vague, graded representation. To do so, we represent the different theoretical boundaries of exurban areas using crisp boundaries and fuzzy membership functions and visualize these empirical boundaries of exurban areas in maps of Delaware County, Ohio, USA using standard GIS techniques.

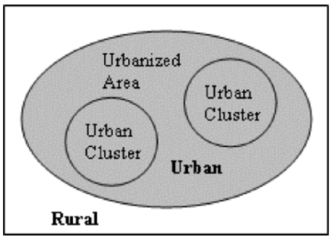

Fig. 1. Idealized spatial configuration of urban and rural area concepts

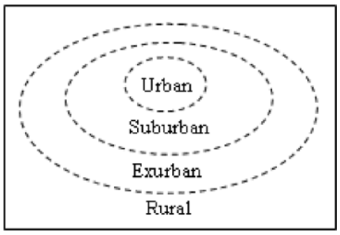

Fig. 2. Simple spatial distribution of urban, suburban, exurban, and rural areas

2 Theoretic background and concept of uncertainty in exurban boundaries

We start by introducing some ontological terms used to identify areas related to an urban environment.

2.1 The concept of urban, suburban, and rural zone

Urban means a zone—also called Urbanized Area—that has a population of at least 50,000 people and a population density of at least 1,000 people per square mile. Such zones are located within an urbanized area or an urban cluster (Fig. 1) [26]. Urban cluster means an area that has census block groups that is contiguous and densely distributed and its census block groups have at least 2,500 people but fewer than 50,000 people. Suburban county means a non-central county classified as metropolitan. Metropolitan counties outside this ring of suburban counties are considered exurban [16].

Rural zone means all areas outside the boundary of an urbanized area. Part of an administrative area—such as a census tract, or county outside metropolitan areas—can belong to urban area(s), and the other part of it can belong to rural area(s) simultaneously [26]. It is hard to say whether an administrative unit is 100% rural or 100% urban. Based on this, we get a simplified spatial distribution of urban, suburban, exurban, and rural zones as Fig. 2. The dotted boundaries between suburban-exurban and exurban-rural indicate that these areas don't have a clear boundary between them. Based on this, we might imagine a conceptual space consist of urban ontology terms. In that space, an axis starts from the term 'urban' and ends at 'rural'. On the axis, the term 'exurban' exist between them but more close to 'rural' with other similar terms such as 'sprawl' or 'periurban'.

Turning our focus on the exurban areas Daniels (1999) define this urban fringe as a region of middle ground between the wide-open rural lands that are beyond commuting distance to a metro area and the expanding suburban residential and commercial developments. Again, as indicated in Fig. 2, it is not easy to define the exact location of exurban areas. The exurban area is also called ‘periurban’ mainly as in the French-speaking literature, ‘deconcentration’, ‘decentralisation’, or ‘extended suburbanisation’, in Europe and North America [4]. According to Caruso (2005), periurbanization refers to the process of residential growth towards the rural periphery of a city. This process leads to the emergence of a spatial zone characterized by a mix of agricultural activities and commuting households [5] (cited in [4]).

2.2 The concept of exurban areas

We mentioned the present lack of consensus on the definition of exurbanization. However, there have been efforts to narrow the uncertainty of the concept. The exurban area could be defined more specifically using for example population and distance from the central city as a basis, although the exact limits are still likely to be different from one researcher to another. For example, according to Daniels (1999), exurban area is 10 to 50 miles away from a major urban center of at least 500,000 people (zone B in Fig. 3), or 5 to 30 miles from a city of at least 50,000 people. This is generally within 25 minute commuting distance and the population density is generally less than 500 people per square mile. Nelson (1992) on the other hand argues for a definition of exurban counties being those within 50 miles of the boundary of the central city of a Metropolitan Statistical Area (MSA) with a population of between 500,000 and less than 2 million (zone A in Fig. 3), or within 70 miles of the boundary of the central city of an MSA with a population of more than 2 million (zone C in Fig. 3). However, Nelson (1992) points out that it is unclear to determine exurbanization when rural areas become more similar to exurban areas, or when exurban areas become more similar to suburban or urban areas.

The existing definitions of exurbanization of both Daniels and Nelson are summarized in Fig. 3.

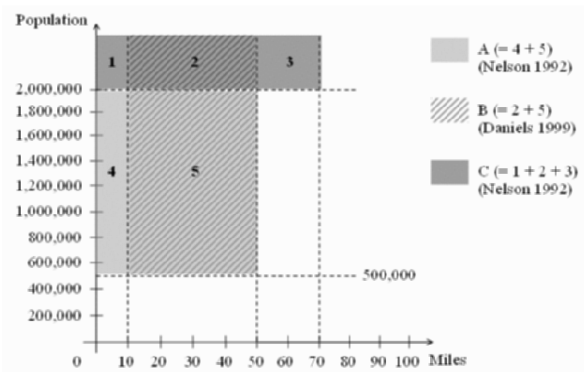

Fig. 3. Differences between existing definitions of exurbanization of Daniels (1999) and Nelson (1992)

In Fig. 3, it is difficult to represent an exact position on the 'Miles' (distance) axis for Nelson's definition since his definition uses the distance from the "boundary of the central city". The extent of central city varies from city to city, and in Fig. 3 we assume the exurban area to be located somewhere between 0 miles and 70 miles but maintain a 50 mile wide band around the city.

These two examples demonstrate that two definitions of exurban areas not only differ in the limits they set for a determining variable such as distance, but also how they use slightly different points of departure for those measurements. It is also apparent that no matter how well defined these urban concepts get in theory, the actual understanding and application is likely to be ambiguous because of their inherently vague character [8]. We therefore propose that these and other terms that will make up future urban ontologies should be defined in a way that explicitly represents their definitional vagueness and makes it possible to evaluate different ontological commitments conceptually as well as spatially. Several approaches have been suggested to represent uncertainty in ontologies. In this work we follow the methodology proposed by Ahlqvist (2005) which is based on the underlying theory of conceptual spaces [10] and that use fuzzy set [31] based formalisms. We formally represent a concept space as a collection, or set, of property definitions. A property definition is represented as a set of values from a certain domain, for example the interval of distance values. The use of fuzzy set based extensions of traditional set theory makes it possible to reconcile the boundaries between different definitions of suburban-exurban and exurban-rural areas acknowledging the graded changes from one zone to another.

3 Exurban areas in Delaware County

Census data is one of the fundamental sources to define the boundary of exurban in terms of calculating total population and population density of unit area. The example dataset consists of block group data of Delaware County, just north of Columbus, Ohio. Population density, the center of Columbus MSA, and urbanized area data all come from U.S. Census2000. Total population of the Columbus MSA was 1,527,948.

The fuzzy membership functions for each definition can be generated by empirical measurement or expert judgment (c.f. [3]). In this case we used the latter approach and elicited membership functions from interpreting exurban areas as a vague concept with the written definitions as a guide to develop membership functions. It is clear that the distance from the center of MSA to each block group increases or decreases linearly. Since the width of transition from one urban zone to another has not been described in the literature yet, we define it arbitrarily for this demonstration. We allow each definition to share fuzzy boundaries with other urban zones. Then we can logically combine multiple memberships, from the distance and population density dimensions, taking the average of the two membership values at any one location and visualize the result in a map.

3.1 Definition of exurban areas according to Daniels (1999)

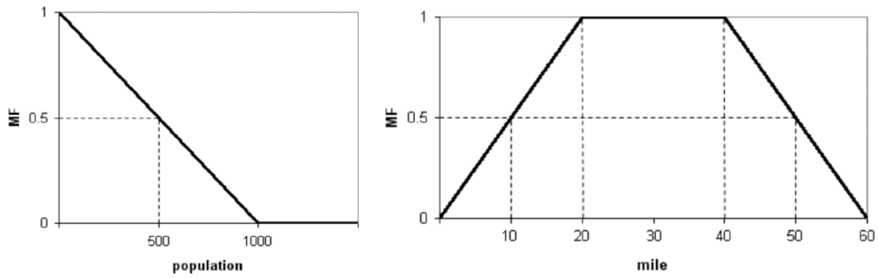

Fig. 4. Membership function of population *(left)* and distance *(right)* in Delaware County based on Daniels's (1999) definition

In Fig. 4 on the left, a simple linear function is used to define the MF value of population. In Fig. 4 on the right, a combination of simple linear functions is assigned to define the MF value of distance. The final membership value for the Daniels definition is determined by calculating the average of the two membership values $MF_{(D)}$ and $MF_{(P)}$.

3.2 Definition of exurban areas according to Nelson (1992)

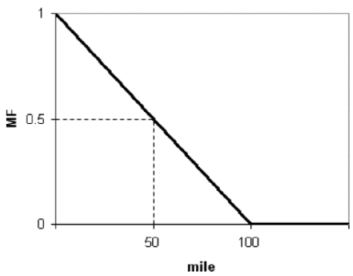

Fig. 5. Membership function of distance in Delaware County based on Nelson's (1992) definition

In Fig. 5, a simple linear function is assigned to define the MF value. In the study area, the whole Delaware County is contained within 50 miles of the boundary of the central city, the City of Columbus. Therefore, the range of distance on the 'Miles' axis for Nelson's definition is assumed to be from 0 to 50 miles in calculating its membership function values in Fig. 5.

4 Results

The maps in Fig. 6 show the difference between crisp membership and fuzzy membership representations of Nelson's (1992) definition of exurban. Since only the distance variable is used in Nelson's area, the map in the left with crisp membership represents the entire area in Delaware County as exurban. However, the map in the right with a fuzzy membership representation shows the gradual transition of MF values of being exurban. From the membership function based on Nelson's definition, areas closer to the center of Columbus MSA have higher MF values of being exurban.

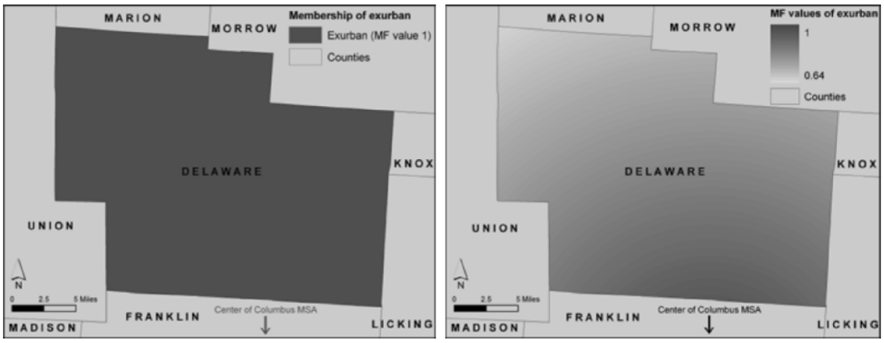

Fig. 6. Exurban areas based on Nelson's (1992) definition with crisp membership *(left)* and fuzzy membership *(right)*

The maps in Fig. 7 show exurban areas with crisp and fuzzy memberships based on Daniels' (1999) definition. In the left map, each block group is assigned to either exurban (MF value 1) or non-exurban (MF value 0). Most of the non-exurban areas are located near central urban areas such as the Columbus MSA in the lower part of the map and near the City of Delaware in the middle-left. In the right map, the block groups have MF values—the degree of being exurban—ranging from 0 to 1 and the MF values are classified into 5 classes for visualization purposes. Since this definition includes distance, the MF values generally increases with distance from the center of MSA. The population component of this definition causes additional variation along the general distance trend. For example, the block groups in the class of values 0.6~0.8 are roughly forming a band around the 10 miles distance from the center of Columbus MSA. Also, some small number of block groups show "leapfrog[1]" pattern in the entire study area. In these two maps, the difference between crisp and fuzzy memberships of the definition is clearly shown. The map with fuzzy membership is able to address a more specific spatial pattern of sprawl in the study area than the map with crisp membership. It shows variations of degree in being exurban among the block groups within the exurban area as well as the non-exurban areas that the map with crisp membership fails to show.

---

[1] If there are discontinuous development pattern s, this is called leapfrog or scattered development (Irwin and Bockstael 2002).

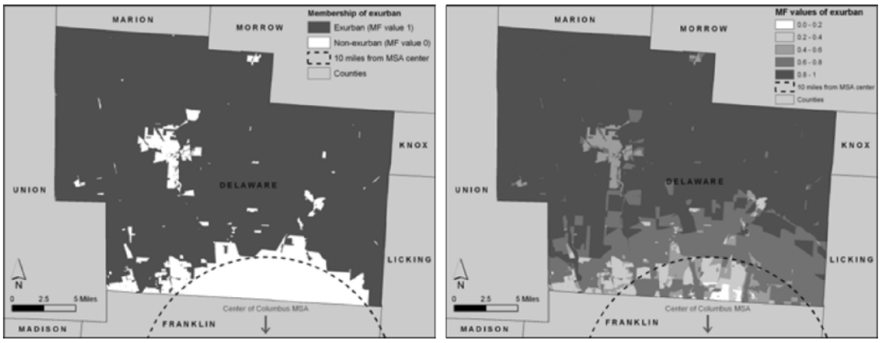

Fig. 7. Exurban areas based on Daniel's (1999) definition with crisp membership *(left)* and fuzzy membership *(right)*

5 Concluding discussion

The relevance of representing exurban areas as vague objects was demonstrated by comparing the traditional crisp representation with a vague, graded representation. The spatial implications of different ontological commitments were compared as they are represented by different definitions of exurban areas. By using the concept of fuzzy ontology in defining exurban boundaries, it is revealed that a clear difference exists between crisp membership and fuzzy membership representations. The crisp classification of exurban area may miss the graded phenomena within such areas.

There are two major differences between our suggested ontology representation and prevailing approaches. First, we propose to generalize the standard first-order logic representation, found in for example the Web Ontology Language (OWL), with a fuzzy-set-based one that can explicitly recognize the vagueness of terms and admit partial belonging to several possible categories. This direction is pointed out by many authors as crucial for developing ontologies to have the expressiveness needed to support practical applications [27];[19]. Although current versions of OWL does not support fuzzy memberships or fuzzy inference per se, we think it is possible to use this or other XML based description languages for fuzzy concept representations, albeit in a not so effective manner. Work on probabilistic and fuzzy extensions [11];[23] of traditional description logics also suggest that it is possible to develop more flexible reasoning capabilities for uncertain ontology semantics. Second, ontology development has this far mostly focused on developing standardized terminologies to support interoperability. This single ontology approach can be contrasted with a hybrid ontology approach [15] in which standardization focuses on the descriptive properties of a term rather than the actual terminology. In our example we can compare different notions of exurban areas by using standard descriptive properties such as population and distance from the city. The combination of these two modifications makes our approach able to compare across heterogeneous terminologies and look for similarities and differences in a flexible manner.

Fisher (1999) argues that ambiguity does come into play in the allocation of social and economic program resources, and it can lead to contention between politicians over the issue of financial support. Urban growth studies of geographic extent, patterns, and classes can be a new resource for local land use decision makers as they plan the future of their communities [29]. In this context, the study of exurbanization showing the uncertainty of its boundaries would also be useful since it reveals the heterogeneity of exurban areas in a location specific context.

This work can be extended in several ways. First of all, we should seek to incorporate the dynamic character of urbanization processes. It is relatively straightforward to visualize the changing exurban boundaries through time in 3D animation of snapshot maps from different points in time. But, we could also further extend the very simplistic category descriptions exemplified here with time dependent characteristics such as time constraints in the definition. This could help identify for example areas where an exurbanization process is just starting. There has been significant work on spatio-temporal modeling using predominantly first-order logic approaches (c.f. [17]). In terms of incorporating dynamic characteristics into our suggested approach, we do not see any major problems to include for example fuzzy time constraints, but this is still a matter of further research. Secondly, a weighted fuzzy membership function can be used for land-use decision making [32]. The subjective weights can lead to large variations [18].

The difference between definitions can be compared by integrating the distribution of fuzzy set values at each location. We might identify satellite nodes within the exurbanization patterns, whose growth is probably related to the central metropolis area and which themselves become secondary sources of exurbanization. Based on these results, such fuzzy membership functions may hence be used to elaborate and test new definitions for exurban areas, in order to better match with the observed empirical phenomenon. We could also investigate the inherent error in the data sets. For example, the Census2000 dataset has confidence intervals in the population density data. By measuring different boundaries using the confidence interval, we can figure out the effect of the error in the data.

References

1. Ahlqvist, O.: Using Uncertain Conceptual Spaces to Translate Between Land Cover Categories. International Journal of Geographical Information Science. 19:7 (2005) 831-857
2. Bishr, Y.: Overcoming the Semantic and Other Barriers to GIS Interoperability. International Journal of Geographic Information Science. 12:4 (1998) 299-314
3. Burrough, P.A., McDonell, R.A.: Principles of Geographical Information Systems. Oxford University Press, New York (1998). Caruso, G.: Integrating Urban Economics and Cellular Automata to model Periurbanisation - Spatial dynamics of residential choice in the presence of neighbourhood externalities. Universitė catholique de Louvain (2005)
5. Cavailhès, J., Peeters, D., Sekeris, E., Thisse, J. F., 2004. The periurban city.why to live between the city and the countryside. Regional Science and Urban Economics 34 (6), 681–703.
6. Daniels, T.: When City and Country Collide. Island Press, Washington, DC. (1999)
7. Fisher, P.: Ch.13 Models of Uncertainty in Spatial Data. In: Longley, P., Goodchild, M.F., Maquire, D.J., Rhind, D.W. (eds.): Geographical Information Systems. Vol. 2. 2nd edn. John Wiley, New York (1999)

8. Fisher, P.: Sorites Paradox and Vague Geographies. Fuzzy Sets and Systems. 113 (2000) 7-18
9. Fonseca, F., Egenhofer, M., Davis, C., Borges, K.: Ontologies and Knowledge Sharing in Urban GIS. CEUS - Computer, Environment and Urban Systems. 24:3 (2000) 232-251
10. Gärdenfors, P.: Conceptual Spaces. MIT Press, Cambridge (2000)
11. Heinsohn, J.: Probabilistic description logics. In: de Mantara, R.L., Pool, D. (eds.): Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence. (1994) 311-318
12. Irwin, E.G., Bockstael, N.E.: Land Use Externalities, Growth Management Policies, and Urban Sprawl. Regional Science and Urban Economics. 34:6 (2004) 705-725
13. Kemp, K., Vckovski, A.: Towards an Ontology of Fields. In: Third International Conference on GeoComputation, Bristol (1998)
14. Leyk, S., Boesch, R., Weibel, R.: A Conceptual Framework for Uncertainty Investigation in Map-based Land Cover Change Modelling. Transactions in GIS. 9:3 (2005) 291-322
15. Lutz, M., Klien, E.: Ontology-based retrieval of geographic information. International Journal of Geographical Information Science. 20:3 (2006) 233-260
16. Nelson, A.C.: Characterizing Exurbia. Journal of Planning Literature. 6:4 (1992) 350-368
17. Peuquet, D.: Representations of Space and Time. Guilford Press, New York (2002)
18. Robinson, V.B.: A Perspective on the Fundamentals of Fuzzy Sets and Their Use in Geographic Information Systems. Transactions in GIS. 7:1 (2003) 3-30
19. Sheth, A., Ramakrishnan, C., Thomas, C.: Semantics for the Semantic Web: The Implicit, the Formal and the Powerful. International Journal on Semantic Web & Information Systems. 1:1 (2005) 1-18
20. Smith, B.: On Drawing Lines on a Map. In Frank, A., Kuhn, W., Mark, D.M. (eds.): Spatial Information Theory - COSIT 95. Vienna (1995) 475-484
21. Smith, B.: An Introduction to Ontology. In: Peuquet, D., Smith, B., Brogaard, B. (eds.): The Ontology of Fields. NCGIA, Bar Harbor (1998) 10-14
22. Smith, B., Mark, D.: Ontology and Geographic Kinds. In: International Symposium on Spatial Data Handling. Vancouver (1998) 308-320
23. Straccia, U.: Reasoning within Fuzzy Description Logics. Journal of Artificial Intelligence Research. 14 (2001) 137-166
24. Teller, J., Keita, A.K., Roussey, C., Laurini, R.: Urban Ontologies for an Improved Communication in Urban Civil Engineering Projects. In: Proceedings of the International Conference on Spatial Analysis and GEOmatics. Research & Developments, SAGEO, Avignon (2005)
25. Theobald, D.M.: Landscape Patterns of Exurban Growth in the USA from 1980 to 2020. Ecology and Society. 10:1 (2005) 32
26. U.S. Census Bureau: Census 2000 Summary File 1 Technical Documentation. U.S. Census Bureau, Washington, D.C. (2001)
27. Uschold, M.: Where are the semantics in the semantic web?. AI Magazine. 24 (2003) 25-36
28. Wiederhold, G., Jannink, J.: Composing Diverse Ontologies. In: 8th Working Conference on Database Semantics (DS-8), Rotorua (1999)
29. Wilson, E.H., Hurd, J.D., Civco, D.L., Prisloe, M.P., Arnold, C.: Development of a Geospatial Model to Quantify, Describe and Map Urban Growth. Remote Sensing of Environment. 86 (2003) 275-285
30. Wolman, H., Galster, G., Hanson, R., Ratcliffe, M., Furdell, K., Sarzymski, A.: The Fundamental Challenge in Measuring Sprawl: Which Land Should be Considered?. The Professional Geographers. 57:1 (2005) 94-105
31. Zadeh, L.A.: Fuzzy sets. Information and Control. 8:3 (1965) 338-353
32. Zeng, T.Q., Zhou, Q.: Optimal Spatial Decision Making Using GIS: A Prototype of a Real Estate Geographical Information System (REGIS). International Journal of Geographical Science. 15 (2001) 307-321