Proceedings of the X Brazilian Symposium on GeoInformatics

Rio de Janeiro, Brazil

December 8-10, 2008

Dados Internacionais de Catalogação na Publicação

681.3.06	Simpósio Brasileiro de Geoinformática (10. , 2008: Rio de Janeiro,RJ).
SI 57	Anais do X Simpósio Brasileiro de Geoinformática, Rio de Janeiro, Brasil, 8 a 10 dez. 2008. Formato Digital ISBN
	1.Geoinformática. I.Carvalho, M. T. M.; Casanova, M. A.; Gattass, M. II.Vinhas, Lúbia. III.Pontícia Universidade Católica do Rio de Janeiro.

Organized by

Promoted by

Sponsored by

Supported by

Preface

The GeoINFO series of scientific symposia is an annual forum for exploring ongoing research, development and innovative applications on geographic information science and related areas. In 2008, to celebrate the 10th anniversary of GeoINFO, the symposium was held again in the City of Rio de Janeiro, where the first GeoINFO took place.

The GeoINFO symposia also invites leading GIScience and spatial database researchers to present to the local community a perspective of the state-of-the-art in the area. In 2008, the guest speakers were Dr. Max Craglia, from the Spatial Data Infrastructures Unit of the Joint Research Centre of the European Commission, and Dr. Stephan Winter, Associate Professor of the Department of Geomatics, University of Melbourne.

Out of the 65 papers that were submitted, the Program Committee members and additional referees carefully selected 10 full and 10 short papers, thus defining an acceptance rate of 30.7%. The authors of the accepted papers represent 15 distinct Brazilian academic institutions and research centers. The collection of the accepted papers indeed express the high quality of the research in the area being carried out in Brazil.

The support of INPE (Brazil's National Institute for Space Research), SBC (Brazilian Computer Society) and Selper (Society of Latin-American Specialists on Remote Sensing) to the GeoINFO series is gratefully acknowledged. Special thanks goes to TecGraf/PUC-Rio for hosting the event in 2008.

The General Chairs and the Program Chair wish to thank specially Terezinha, Hilcéa, Daniela and Janete, from INPE, and Ruth and Sandra, from PUC, for their companionship in this journey.

Marcelo Tílio Monteiro de Carvalho
General chair

Marcelo Gattass
General chair

Marco Antonio Casanova
General chair

Lúbia Vinhas
Program chair

Conference Organization

General Chairs

Marcelo Tílio Monteiro de Carvalho
TecGraf, Dept. Informatics, PUC-Rio, Brazil

Marco Antonio Casanova
Dept. Informatics, PUC-Rio, Brazil

Marcelo Gattass
Dept. Informatics, PUC-Rio, Brazil

Program Chair

Lúbia Vinhas
INPE, Brazil

Local Organization

Daniela Seki
INPE

Hilcéa Santos Ferreira
INPE

Janete da Cunha
INPE

Ruth Maria Barros Fagundes de Sousa
PUC-Rio

Sandra Lúcia Barbosa Schwabe
PUC-Rio

Terezinha Gomes dos Santos
INPE

Program committee

Lúbia Vinhas, INPE/DPI
Marcelo Tílio Monteiro de Carvalho, Tecgraf/PUC-Rio
Ana Lúcia Bezerra Candeias, UFPE
Andrea Rodriguez, Universidad de Concepción
Antonio Carlos Rocha Costa, UCPel
Antonio Machado, Kepler Sistemas
Antonio Miguel Vieira Monteiro, INPE
Armanda Rodrigues, Universidade Nova de Lisboa
Camilo Rennó, INPE
Claudio Baptista, UFCG
Claudio Esperança, COPPE/UFRJ
Clodoveu Davis, PUC Minas
Chris Jones, Cardiff University
Christelle Vangenot, Ecole P. Fédérale Lausanne
Dan Brown, University of Michigan
Eduardo Camargo, INPE
Eymar Lopes, INPE
Frederico Fonseca, Penn State
Frederico Liporace, Kepler Sistemas
Guaraci Erthal, INPE
Gilberto Câmara, INPE
Holger Schwarz, University of Stuttgart
João Argemiro Paiva, Oracle Coorporation
Jorge Campos, UNIFACS
José Alberto Quintanilha, USP
Julio D'Alge, INPE
Jussara Ortiz, INPE
Karla Borges, PRODABEL
Laercio Namikawa, INPE
Leila Maria Garcia Fonseca, INPE
Leonardo Bins, INPE
Marcelino Pereira, UERN
Marcelo Azevedo, UFMG
Marco Antonio Casanova, PUC-Rio
Maria Isabel Sobral Escada, INPE
Margareth Simões, UERJ
Mário Gaspar da Silva, Universidade de Lisboa
Max Egenhofer, University of Maine
Miguel Torres, Centro de Investigacion en Computacion
Paulo Graça, INPA
Raul Feitosa, PUC-Rio
Ricardo Ciferri, UFSCAR
Ricardo Torres, UNICAMP
Sergio Faria, UFMG
Sergio Rosim, INPE
Shashi Shekkar, University of Minnesota
Silvana Amaral, INPE
Stephen Winter, University of Melbourne
Tiago Carneiro, UFOP
Valeria Soares, UFPB
Vania Bogorny, UFRGS
Virginia Ragoni, INPE

External Referees

Demetrius Nunes
Lin Tzy Li
Luciano Digiampietri
Marcelo Metello
Melissa Lemos
Milton Jonathan
Thiago Luís Lopes Siqueira

Contents

Investigating the Effects of Spatial Data Redundancy in Query Performance over Geographical Data Warehouses[1]

Thiago Luís Lopes Siqueira[1], Ricardo Rodrigues Ciferri[1], Valéria Cesário Times[2], Cristina Dutra de Aguiar Ciferri[3]

[1]Departamento de Computação – Universidade Federal de São Carlos (UFSCar)
Caixa Postal 676 – 13.565-905 – São Carlos, SP – Brasil

[2]Centro de Informática – Universidade Federal de Pernambuco – Recife, PE – Brasil

[3]Departamento de Ciências de Computação – Universidade de São Paulo
São Carlos, SP – Brasil

{thiago_siqueira,ricardo}@dc.ufscar.br,
vct@cin.ufpe.br, cdac@icmc.usp.br

Abstract. *Geographical Data Warehouses (GDW) are one of the main technologies used in decision-making processes and spatial analysis. For these, several conceptual and logical data models have been proposed in the literature. However, little attention has been devoted to the study of how spatial data redundancy affects query performance over GDW. In this paper, we investigate this issue. Firstly, we compare redundant and non-redundant GDW schemas and conclude that redundancy is related to high performance losses. Further, we analyze the indexing issue, aiming at improving query performance on a redundant GDW. Comparisons among the SB-index approach, the star-join aided by R-tree and the star-join aided by GiST showed that SB-index significantly improves the elapsed time on query processing from 25% up to 95%.*

1. Introduction

Although Geographical Information System (GIS), Data Warehouse (DW) and On-Line Analytical Processing (OLAP) have different purposes, all of them converge in one aspect: decision-making support. Some authors have already proposed their integration, by making use of a Geographical Data Warehouse (GDW) to provide a means of carrying out spatial analyses together with agile and flexible multidimensional analytical queries over huge data volumes [Fidalgo et al. 2004; Malinowski and Zimányi 2004; Silva et al. 2008; Sampaio et al. 2006]. However, little attention has been devoted to the investigation of the following issue: **how does the spatial data redundancy affect query response time and storage requirements in a GDW**?

Similarly to a GIS, a GDW-based system manipulates geographical data with geometric and descriptive attributes and supports spatial analyses and ad-hoc query windows as well. Like a DW [Kimball and Ross 2007], a GDW is a subject-oriented,

[1] This work has been supported by the following Brazilian research agencies: CAPES, CNPq, FAPESP, FINEP and INEP. The first two authors also thank the support of the Web-PIDE Project in the context of the Observatory of the Education of the Brazilian Government.

integrated, time-variant and non-volatile dimensional database, that is often organized in a star schema with dimension and fact tables. A dimension table contains descriptive data and is typically designed in hierarchical structures in order to support different levels of aggregation. A fact table addresses numerical, additive and continuously valued measuring data, and maintains dimensions keys at their lower granularity level and one or more numeric measures as well. In fact, a GDW stores geographical data into one or more dimensions or into at least one measure [Stefanovic et al. 2000; Fidalgo et al. 2004; Malinowski and Zimányi 2004; Sampaio et al. 2006; Silva et al. 2008]. OLAP is the technology that provides strategic multidimensional queries over the DW. Further, Spatial OLAP (SOLAP) tools provide analytical multidimensional queries based on spatial predicates that mostly run over GDW [Gaede and Günther 1998; Kimball and Ross 2002; Fidalgo et al. 2004; Silva et al. 2008]. Often, the spatial query window is an ad-hoc area, not predefined in the spatial dimension tables.

Attribute hierarchies in dimension tables lead to an intrinsically redundant schema, which provides a means of executing roll-up and drill-down operations. These OLAP operations perform data aggregation and disaggregation, according to higher and lower granularity levels, respectively. Kimball and Ross (2002) stated that a redundant DW schema is preferable to a normalized one, since the former does not introduce new joins costs and the attributes assume conventional data types that require only a few bytes. On the other hand, in GDW it may not be feasible to estimate the storage requirements for a spatial object represented by a set of geometric data [Stefanovic et al. 2000]. Also, evaluating a spatial predicate is much more expensive than executing a conventional one [Gaede and Günther 1998].

Hence, choosing between a normalized and a redundant GDW schema may not lead to the same option as for conventional DW. Then, an experimental evaluation approach should aid GDW designers to make this choice. In this paper we investigate spatial data redundancy effects in query performance over GDW. We mainly analyze if a redundant schema aids SOLAP query processing in a GDW, as it does for OLAP and DW. We also address the indexing issue aiming at improving query processing on a redundant GDW.

This paper is organized as follows. Section 2 discusses related work, Section 3 describes the experimental setup, and Section 4 shows performance results for redundant and non-redundant GDW schemas, using database systems resources. Section 5 describes indices for DW and GDW, and the SB-index structure that we proposed in Siqueira et al. (2009). Section 6 details the experiments regarding the SB-index for the redundant and non-redundant GDW schemas. Section 7 concludes the paper.

2. Related Work

Stefanovic et al. (2000) were the first to propose a GDW framework, which addresses spatial dimensions and spatial measures. Particularly, in a spatial-to-spatial dimension table, all levels of an attribute hierarchy maintain geometric features representing spatial objects. However, the authors do not discuss the effects of such redundant schema, by just focusing on the selective materialization of spatial measures. Fidalgo et al. (2004) foresaw that spatial data redundancy might deteriorate GDW performance. Because of this, they proposed a framework for designing geographical dimensional schemas that strictly avoid spatial data redundancy. The authors also validated their proposal adapting a DW schema for a GDW. Even so, this validation work does not compare if the normalized schema performs SOLAP queries better than redundant ones. Finally, although Sampaio et al. (2006) have proposed a logical multidimensional model, they do not discussed about spatial data redundancy, but reuse the mentioned spatial-to-spatial dimension tables. Therefore, as far as we know, none of the related work outlined in this section have experimentally investigated the effects of spatial data redundancy in GDW, nor examined if this issue really affects SOLAP queries performance. This experimental evaluation is the main objective of our work.

3. Experimental Setup

3.1. Workbench

Experiments were conducted on a computer with 2.8 GHz Pentium D processor, 2 GB of main memory, 7200 RPM SATA 320 GB hard disk, Linux CentOS 5.2, PostgreSQL 8.2.5 and PostGIS 1.3.3. We have chosen the PostgreSQL/PostGIS DBMS because it is a prominent open source database management system. We also adapted the Star Schema Benchmark (SSB) [O'Neil, P., O'Neil, E. and Chen 2007] in order to support GDW spatial analysis, once its spatial information is strictly textual and maintained into Supplier and Customer dimensions. SSB is derived from TPC-H (www.tpc.org/tpch). The changes we made preserved descriptive data and created an spatial hierarchy based on the previously defined conventional dimensions. Both Supplier and Customer dimension tables have the following hierarchy: region $\prec$ nation $\prec$ city $\prec$ address. While the domains of the attributes *s_address* and *c_address* are disjoint, suppliers and customers share city, nation and region locations as attributes.

The following two GDW schemas were developed in order to investigate how much SOLAP queries performance is affected by spatial data redundancy. According to Stefanovic et al. (2000), Supplier and Customer are spatial-to-spatial dimensions and may maintain geographical data as shown in Figure 1: attributes with suffix "*_geo*" store the geometric data of geographical features, and there is spatial data redundancy. For instance, a map for Brazil is stored in every row whose supplier is located in Brazil. However, as stated by Fidalgo et al. (2004), in a GDW, spatial data must not be redundant and should be shared whenever is possible. Thus, Figure 2 illustrates Customer and Supplier sharing city, nation and region locations, but not addresses.

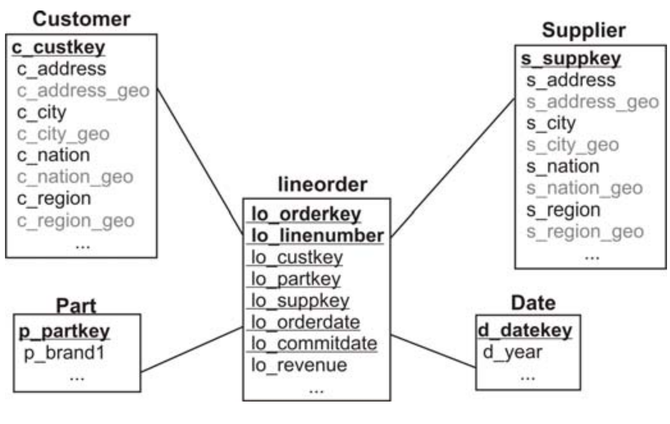


Figure 1. SSB adapted with spatial data redundancy: GRSSB.

While the schema shown in Figure 1 is named as GRSSB (Geographical Redundant SSB), the schema given in Figure 2 is called GHSSB (Geographical Hybrid SSB). Both schemas were created using SSB's database with scale factor 10. Data generation produced 60 million tuples in the fact table, 5 distinct regions, 25 nations per region, 10 cities per nation and a certain quantity of addresses per city that varies from 349 to 455. Cities, nations and regions were represented by polygons, while addresses were expressed by points. All geometries were adapted from Tiger/Line (www.census.gov/geo/www/tiger). GRSSB occupies 150 GB while GHSSB has 15 GB.

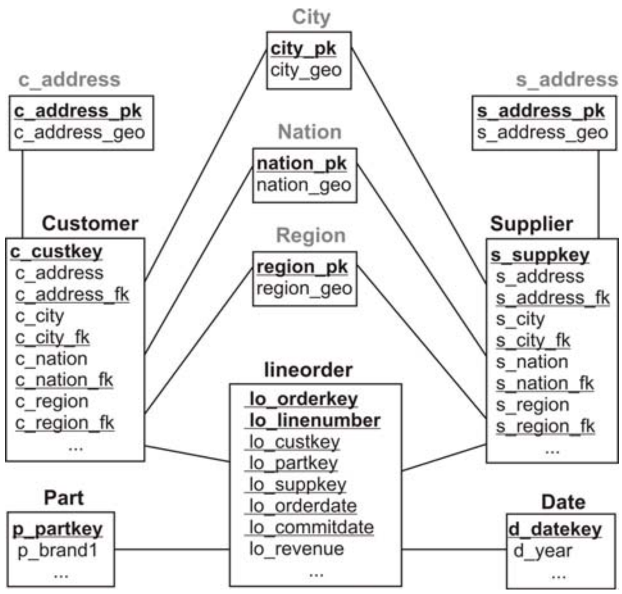


Figure 2. SSB adapted without spatial data redundancy: GHSSB.

3.2. Workload

The workload was based on Query Q2.3 from SSB, showed on Figure 3. We replaced the underlined conventional predicate with spatial predicates involving ad-hoc query windows (QW). These quadratic windows have had a correlated distribution with spatial data, and their centroids are random supplier addresses. Their sizes are proportional to the spatial granularity. Addresses were evaluated with containment range query [Gaede and Günther 1998] and its QW covers 0.001% of the extent. City, nation and region levels were evaluated with intersection range query [Gaede and Günther 1998] and their QW cover 0.05%, 0.1% and 1% of the extent, respectively.

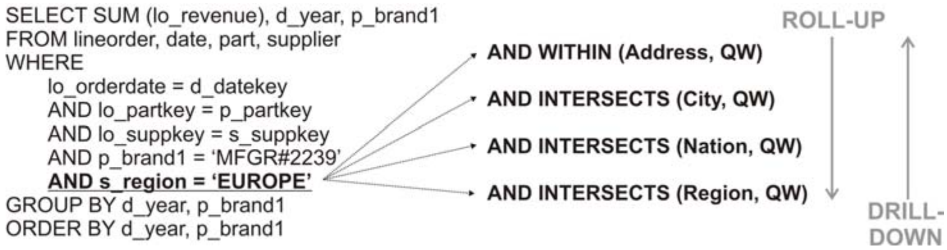


Figure 3. Adaption of SSB Q2.3 query.

The cited query windows allow the aggregation of data in different spatial granularity levels, i.e. the application of spatial roll-up and drill-down operations. Figure 3 also shows how complete spatial roll-up and drill-down operations are performed. Both operations comprise four different size query windows that share the same centroid. Space limitations lead us to post further details about data and query distribution at http://gbd.dc.ufscar.br/papers/geoinfo08/figures.pdf.

4. Performance Results and Evaluation Applied for Different GDW Schemas

In this Section, we discuss the performance results related to the following two test configurations based on the GRSSB and GHSSB schemas: (C1) star-join computation aided by R-tree index [Guttman 1984] on spatial attributes; (C2) star-join computation aided by GiST index (http://gist.cs.berkeley.edu) on spatial attributes. The configurations reflect current available DBMS resources for GDW query processing. Experiments were performed by submitting 5 complete roll-up operations to GRSSB and GHSSB and by taking the average of the measurements for each granularity level. To analyze query processing we gathered the elapsed time (seconds). Table 1 shows these results.

Table 1. Performance results for configurations C1 and C2.

Query processing: elapsed time

	GRSSB C1	GHSSB C1	GRSSB C2	GHSSB C2
Address	2854.17	2866.51	2831.23	2853.85
City	2773.39	2763.17	2773.10	2758.70
Nation	4047.35	2766.14	3449.76	2765.61
Region	6220.68	2787.83	6200.44	2790.29

Fidalgo et al. (2004) foresaw that, although a normalized GDW schema introduces new join costs, it should require less storage space than a redundant one. Our quantitative experiments verified this fact and also showed that SOLAP query processing is negatively affected by spatial data redundancy in a GDW. According to our results, attribute granularity does not affect query processing in GHSSB. For instance, Region-level queries took 2787.83 seconds in C1, while Address-level queries took 2866.51 seconds for the same configuration, a very small increase of 2.82%. Therefore, our experiments showed that star-join is the most costly process in GHSSB. On the other hand, as the granularity level increases, the longer it takes to process a query in GRSSB. For example, in C2, a query on the finest granularity level (Address) took 2831.23 seconds, while on the highest granularity level (Region) it took 6200.44 seconds, an increase of 119%, despite the existence of efficient indices on spatial attributes. Therefore, it is possible to conclude that spatial data redundancy from GRSSB caused larger performance losses to SOLAP query processing, as compared to losses caused by additional joins in GHSSB.

Although our tests have pointed out that SOLAP queries processing in GHSSB is more efficient than in GRSSB, they also showed that both schemas offer prohibitive query response time. Clearly, the challenge in GDW is to retrieve data related to ad-hoc query windows and to avoid the high cost of star-joins. Thus, mechanisms to provide efficient query processing, as index structures, are essential. In the next sections, we describe existing indices and resume a novel index structure for GDW called SB-index [Siqueira et al. 2009], as well as discuss how spatial data redundancy affects them.

5. Indices

In the previous sections we identified some reasons for using an index structure to improve SOLAP query performance over GDW. Section 5.1 discusses the existing indices for DW and GDW, while in Section 5.2, we describe our SB-index approach.

5.1. Indices for DW and GDW

The Projection Index and the Bitmap Index were proven to be valuable choices in conventional DW indexing, since the multidimensionality is not an obstacle to them [O'Neil, P. and Quass 1997; Wu et al. 2006; Stockinger and Wu 2007; Wu et al. 2008]. Consider **X** as an attribute of a relation **R**. Then, the Projection Index on **X** is a sequence of values for **X** extracted from **R** and sorted by the row number. Surely, repeated values may exist. The basic Bitmap index associates one bit-vector to each distinct value ***v*** of the indexed attribute **X**. The bit-vectors maintain as many bits as the number of records found in the data set. If for the k-th record of the data set, we have that **X** = ***v***, then the k-th bit of the bit-vector associated with ***v*** has the value of one. The attribute cardinality, |**X**|, is the number of distinct values of **X** and determines the number of bit-vectors. Also, a star-join Bitmap index can be created for the attribute **C** of a dimension table, in order to indicate the set of rows in the fact table to be joined with a certain value of **C**.

Figure 4 shows data from a DW star-schema, a projection index and bit-vectors of two star-join Bitmap indices. The attribute s_suppkey is referenced by lo_suppkey shown in the fact table. Although s_address is not involved in the star join, it is possible to index this attribute by a star-join Bitmap index, since there is a 1:1 relationship between s_suppkey and s_address. $Q_1 \prec Q_2$ if, and only if it is possible to answer Q_1 using just the results of Q_2, and $Q_1 \neq Q_2$ [Harinarayan et al. 1996]. Therefore, it is possible to index s_region because s_region $\prec$ s_nation $\prec$ s_city $\prec$ s_address. For instance, executing a bitwise OR with the bit-vectors for s_address = B and s_address = C results in the bit-vector for s_region = 'EUROPE'. Visibly, creating star-join Bitmap indices to attribute hierarchies is enough to allow roll-up and drill-down operations.

DIMENSION TABLE: Supplier

s_suppkey	s_address	s_city	s_nation	s_region
1	A	VIETNAM 2	VIETNAM	ASIA
2	B	FRANCE 5	FRANCE	EUROPE
3	C	ROMANIA 2	ROMANIA	EUROPE
4	D	ALGERIA 6	ALGERIA	AFRICA
5	E	ALGERIA 0	ALGERIA	AFRICA

Projection Index

s_nation
VIETNAM
FRANCE
ROMANIA
ALGERIA
ALGERIA

FACT TABLE: Lineorder

lo_suppkey	lo_custkey	rev
1	235	20
1	512	16
2	512	22
3	235	19
3	512	15
3	106	21
4	235	20
5	106	18

Star-join Bitmap: s_address

A	B	C	D	E
1	0	0	0	0
1	0	0	0	0
0	1	0	0	0
0	0	1	0	0
0	0	1	0	0
0	0	1	0	0
0	0	0	1	0
0	0	0	0	1

Star-join Bitmap: s_region

ASIA	EUROPE	AFRICA
1	0	0
1	0	0
0	1	0
0	1	0
0	1	0
0	1	0
0	0	1
0	0	1

Figure 4. Fragment of data, Projection and Bitmap indices.

Figure 4 also illustrates that instead of storing the value 'EUROPE' twice, the Bitmap Index associates this value to a bit-vector and uses 2 bits to indicate the 2 tuples where s_region = 'EUROPE'. This simple example shows how Bitmap deals with data redundancy. Moreover, if a query asking for s_region = 'EUROPE' and lo_custkey = 235 had been submitted for execution, then a bit-wise AND operation is executed using the corresponding bit-vectors, in order to answer this query. Such property justifies why the number of dimension tables does not drastically affect the Bitmap efficiency.

High cardinality has been seen as the Bitmap's main drawback. However, recent studies have demonstrated that, even with very high cardinality, the Bitmap index approach can provide an acceptable response time and storage utilization [Wu et al. 2006, 2008]. Three techniques have been proposed to reduce the Bitmap index size [Stockinger and Wu 2007]: compression, binning and encoding. FastBit is a Free Software that efficiently implements Bitmap indices with encoding, binning and compression [Wu et al. 2006; O'Neil, E., O'Neil, P., and Wu 2007]. FastBit creates a Projection Index and an index file for each attribute. The index file stores one array of distinct values in ascending order, whose entries point to bit-vectors.

Although being suitable for conventional DW, to the best of our knowledge, the Bitmap approach has never been applied to GDW, nor FastBit has resources to support GDW indexing. In fact, just a single GDW index has so far been found in the database literature, namely the aR-Tree [Papadias et al. 2001]. The aR-tree index uses the R-tree's partitioning method to create **implicit ad-hoc hierarchies for the spatial objects**. Nevertheless, this is not the case for some GDW applications which mainly make use of predefined attribute hierarchies, such as region $\prec$ nation $\prec$ city $\prec$ address. Thus, there is a lack of an efficient index for GDW to support predefined spatial attribute hierarchies and to deal with multidimensionality.

5.2. The SB-index

The Spatial Bitmap Index (SB-index) aims at introducing the Bitmap index in GDW in order to reuse this method's legacy for DW and inherit all of its advantages. SB-index computes the spatial predicate and transforms it into a conventional one, which can be computed together with other conventional predicates. This strategy provides the whole query answer using a star-join Bitmap index. As a result, the GDW star-join operation is avoided and **predefined spatial attributes hierarchies can be indexed**. Therefore, the SB-index provides a means of processing SOLAP operations.

We have designed the SB-index as an adapted projection index. Each entry maintains a key value and a minimum bounding rectangle (MBR). The key value references the spatial dimension table's primary key and also identifies the spatial object represented by the corresponding MBR. DW often have surrogate keys, and thus, an integer value is an appropriate choice for primary keys. A MBR is implemented as two pairs of coordinates. As a result, the size of SB-index entry (denoted as **s**) is given by the expression **s** = sizeof(int) + 4* sizeof(double).

SB-index is an array stored on disk. L is the maximum number of SB-index objects that can be stored into one disk page. Suppose that the size **s** of an entry is equal to 36 bytes (one integer of 4 bytes and four double precision numbers of 8 bytes each) and that a disk page has 4 KB. Thus, we have L = (4096 DIV 36) = 113 entries. In order to avoid fragmented entries among different pages, some unused bytes (U) may separate them. In this example, U = 4096 – (113×36) = 28 bytes.

The SB-index query processing, illustrated in Figure 5, has been divided into two tasks. The first one performs a sequential scan on the SB-index and of adding candidates (possible answers) to a collection. Then, the second task consists of checking which candidates are seen as answers and of producing a conventional predicate.

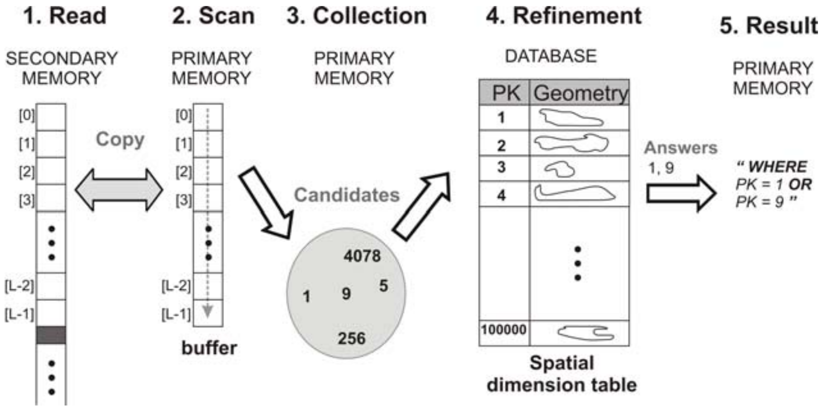


Figure 5. The SB-index Query Processing.

During the scan on the SB-index file, one disk page must be read per step, and its contents copied to a buffer in primary memory. After fulfilling the buffer, a sequential scan on it is needed to test, for each entry, the MBR against the spatial predicate. If this test evaluates to true, the corresponding key value must be collected. After scanning the whole file, a collection of candidates is found (i.e., key values whose spatial objects may be answers to the query spatial predicate). Currently, the SB-index supports intersection, containment and enclosure range queries, as well as point and exact queries.

The second task must check which candidates can really be seen as answers. This is the refinement task and requires an access to the spatial dimension table using each candidate key value, in order to fetch the original spatial object. Clearly, indicating candidates reduces the number of objects that must be queried, and is a good practice to decrease query response time. Then, the previously identified objects are tested against the spatial predicate using proper spatial database functions. If this test evaluates to true, the corresponding key value must be used to compose a conventional predicate. Here the SB-index transforms the spatial predicate into a conventional one. Finally, the resulting conventional predicate is passed to the FastBit software, which can combine it with other conventional predicates and resolve the entire query.

6. Performance Results and Evaluation Applied to Indices

In this Section, we discuss the performance results derived from the following additional test configuration: (C3) SB-index to avoid a star-join. This configuration was applied to the same data and queries used for configurations C1 and C2, and aims at comparing both C1 and C2 (given in Section 3) to C3. We reused the Experimental Setup outlined in Section 3. Our index structure was implemented with C++ and compiled with gcc 4.1. Disk page size was set to 4 KB. Bitmap indices were built with WAH compression [Wu et al. 2006] and without encoding or binning methods. To investigate the index building costs, we gathered: quantity of disk accesses, elapsed time (seconds) and space required to store the index. To analyze query processing we collected the elapsed time (seconds) and the number of disk accesses. In the experiments, we submitted 5 complete roll-up operations to schemas GRSSB and GHSSB and took the average of the measurements for each granularity level.

6.1. First Battery of Tests

Table 2 shows the results derived from applying the SB-indices building operations to GRSSB and GHSSB. Due to spatial data redundancy, in GRSSB all indices have 100,000 entries (consequently the same size) and require the same number of disk accesses to be built. In addition, one spatial object has its MBR obtained several times, since this object may be found in different tuples, causing an overhead. The star-join Bitmap index occupies 2.3 GB, and took 11,237.70 seconds to be built. Each SB-index requires 3.5 MB, i.e., only a small addition to storage requirements (0.14% for each granularity level).

Table 2. Measurements on building SB-index for GRSSB and GHSSB.

	SB-indices building: GRSSB			SB-indices building: GHSSB		
	Elapsed time	Disk accesses	Disk utilization	Elapsed time	Disk accesses	Disk utilization
Address	48	886	3.5 MB	18	886	3.5 MB
City	1,856	886	3.5 MB	4	4	16 KB
Nation	11,566	886	3.5 MB	4	2	8 KB
Region	19,453	886	3.5 MB	2	2	8 KB

Due to normalization, in GHSSB, each spatial object is stored once in each spatial dimension table. Therefore, disk accesses, elapsed time and disk utilization assume lower values than that mentioned for GRSSB, during SB-index building. The star-join Bitmap index occupies 3.4 GB, and took 12,437.70 seconds to be built. SB-index adds at most 0.10% to storage requirements, specifically at the Address level.

Table 3 shows the SB-index query processing results. The time reduction columns compare how much faster C3 is than the best result between C1 and C2 (Table 1). SB-index can check if a point (address) is within a query window without refinement. However, to check if a polygon (city, nation or region) intersects the query window, a refinement task is necessary. Thus, Address level has a better performance gain.

Query processing using SB-index in GRSSB performed 886 disk accesses independently from the chosen attribute. This fixed number of disk accesses is due to the fact that we used sequential scan and all indices have the same number of entries for every granularity (i.e., all indices have the same size). Redundancy affected negatively SB-index, since the same MBR may be evaluated several times during both the SB-index sequential scan and the refinement task. On the other hand, adding only 0.14% to space requirements causes a query response time reduction between 25% and 95%, justifying the adoption of SB-index in GRSSB. In GHSSB, time reduction was always superior to 90%, even better than in GRSSB. This difference lies on the fact that indices had become smaller, due to spatial data redundancy avoidance. In addition, each MBR is evaluated only once, in contrast to GRSSB. Again, the tiny SB-index requires little storage space and contributes to reduce query response time drastically.

Table 3. Measurements on query processing with SB-index.

	Query processing: GRSSB			Query processing: GHSSB			
	C3 Elapsed Time	C3 Time reduction	C3 Disk accesses	C3 Elapsed Time	C3 Time reduction	C3 Disk accesses	
Address	132.41	95.32%	886	131.91	95.38%	886	**Address**
City	271.74	90.20%	886	150.00	94.56%	4	**City**
Nation	1178.12	65.84%	886	201.70	92.70%	2	**Nation**
Region	4621.99	25.45%	886	268.37	90.38%	2	**Region**

6.2 Redundancy Enhancement

As shown in Section 6.1, the motivation behind carrying out the simple and efficient SB-index modification is to be able to efficiently deal with spatial data redundancy. An overhead is caused while querying GRSSB because repeated MBR are evaluated and repeated spatial objects are checked in the refinement phase. On the other hand, unique MBR values should eliminate this overhead, as is the GHSSB case. Therefore, we proposed a second level to SB-index, which consists of assigning a list for each distinct MBR of all key values associated with it. The lists are stored on disk. On query processing, each distinct MBR must be tested against the spatial predicate. If this comparison evaluates to true, one key value from the corresponding list immediately guides the refinement. If the current spatial object is an answer, then all key values from the list are instantly added to the conventional predicate. Finally, the conventional predicate is passed to FastBit, which solves the entire query. As shown in Table 4, the modification severely decreased the number of disk accesses and the query response time in GRSSB. Address granularity level was not tested because it is not redundant. The adaption requires a very small fraction of the star-join Bitmap index volume: from 0.16% to 0.20% for City and Region granularity levels, respectively. Also, time reduction varies from 80.40% to 91.73%.

Table 4. Measurements on the enhancement of SB-index applied to GRSSB.

	Indices building			Query processing			
	Elapsed time	Disk accesses	Disk utilization	Disk accesses	Elapsed time	Time reduction	
City	2,005	101,385	4.81 MB	511	229.11	91.73%	**City**
Nation	11,428	100,935	3.93 MB	57	507.41	85.29%	**Nation**
Region	19,446	100,897	3.86 MB	36	1214.93	80.40%	**Region**

Finally, Figures 6 and 7 respectively shows how SB-index performed in GRSSB and GHSSB, by comparing it to the best result between C1 and C2. The axes indicate the time reduction provided by SB-index. In Figure 6, except for granularity level Address, the results are from SB-index second level enhancement. Instead of suggesting the elimination of the redundancy on GDW schemas, we proposed a means of reducing its effects. This strategy enhances SB-index portability, since it allows using the SB-index in distinct GDW schemas.

7. Conclusions

This paper has analyzed the effects of spatial data redundancy in query performance over Geographical Data Warehouses (GDW). In order to carry out such analysis, we compared the query response times of spatial roll-up operations for two distinct GDW schemas. Since redundancy is related to attribute hierarchies in dimension tables, the first schema, GRSSB, was designed intrinsically redundant, while the other, GHSSB, avoids redundancy through normalization. Our performance tests, using current database management systems resources, showed that GRSSB's spatial data redundancy introduced more performance losses than GHSSB's joins costs.

These results motivated us to investigate indexing alternatives aiming at improving query performance on a redundant GDW. Comparisons among the SB-index, and the star-join aided by efficient spatial indices (R-trees and GiST) showed that SB-index greatly improved query processing: in GRSSB, performance gains were from 25.45% to 95.32%, while in GHSSB were from 90.38% to 95.38%. The results also showed that SB-index is very compact compared to the star-join Bitmap, requiring a very small fraction of this index volume: 0.10% to 0.14%. The lower improvement obtained in GRSSB motivated us to propose a specific enhancement on SB-index to deal with spatial data redundancy, by evaluating distinct MBR and spatial objects once, instead of multiple times. This enhancement provided performance gains from 80.40% to 91.73% in GRSSB. Clearly, spatial data redundancy is related to performance losses.

We are currently investigating some other strategies for minimizing the data redundancy effects on the SB-index, such as adapting a SB-index to use the R*-tree CR to manipulate distinct MBR [Beckmann et al. 1990; Gaede and Günther 1998]. In order to complement the investigation of the data redundancy effects, we are planning to run new experiments using increasing data volumes and different database management systems.

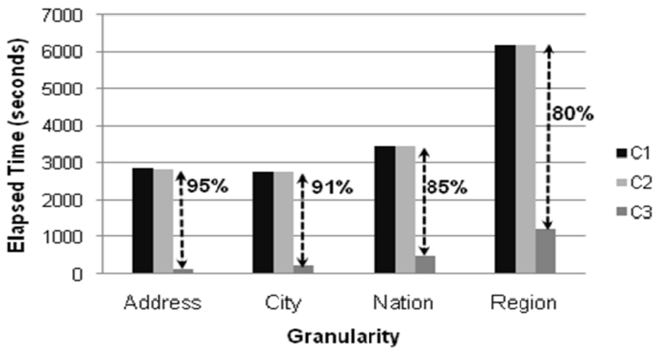


Figure 6. How SB-index performed better than C1 and C2 in GRSSB.

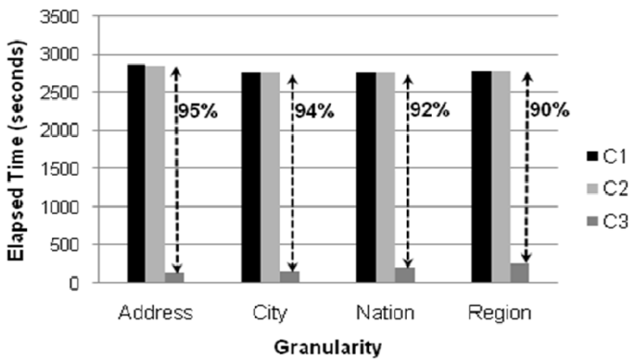


Figure 7. How SB-index performed better than C1 and C2 in GHSSB.

References

Beckmann, N., Kriegel, H.-P., Schneider, R., Seeger, B. (1990) "The R*-Tree: An Efficient and Robust Access Method for Points and Rectangles". In: SIGMOD Conference. p.322-331.

Fidalgo, R. N. et al. (2004) GeoDWFrame: "A Framework for Guiding the Design of Geographical Dimensional Schemas". In: 6th DaWak. p. 26-37.

Gaede, V., Günther, O. (1998) "Multidimensional Access Methods". In *ACM Computing Surveys*, v.30, n.2, p.170-231.

Guttman, A. (1984) "R-Trees: A Dynamic Index Structure for Spatial Searching". In *ACM SIGMOD Record*, v.14, n.2, p.47-57.

Harinarayan, V., Rajaraman, A. and Ullman, J. D. (1996) "Implementing Data Cubes Efficiently". In *ACM SIGMOD Record*, v.25, n.2, p.205-216.

Kimball, R. and Ross, M. (2002) The Data Warehouse Toolkit. Wiley, 2nd edition.

Malinowski, E. and Zimányi, E. (2004) "Representing Spatiality in a Conceptual Multidimensional Model". In: 12th ACM GIS. p.12-22.

O'Neil, E., O'Neil, P., and Wu, K. (2007) "Bitmap Index Design Choices and Their Performance Implications". In: 11th IEEE IDEAS. p. 72-84.

O'Neil, P., and Graefe, G. (1995) "Multi-Table Joins Through Bitmapped Join Indices". In *ACM SIGMOD Record*, v.24, n.3, p.8-11.

O'Neil, P., O'Neil, E. and Chen, X. (2007) "The Star Schema Benchmark", http://www.cs.umb.edu/~poneil/starschemab.pdf. July.

O'Neil, P. and Quass, D. (1997) Improved Query Performance with Variant Indexes, In: ACM SIGMOD Conference. p.38-49.

Papadias, D. et al. (2001) "Efficient OLAP Operations in Spatial Data Warehouses". In: 7th Symposium on Spatial and Temporal Databases. p. 443-459.

Sampaio, M. C., et al. (2006) "Towards a logical multidimensional model for spatial data warehousing and OLAP". In: 8th ACM DOLAP, p. 83-90.

Silva, J., Oliveira, A., Salgado, A. C., Times, V. C., Fidalgo, R., Souza C. (2008) "A set of aggregation functions for spatial measures". In 11th ACM DOLAP.

Siqueira, T. L. L., Ciferri, R. R., Times, V. C., Ciferri, C. D. A (2009) "An Spatial Bitmap-Based Index for Geographical Data Warehouses". In: 24th ACM SAC.

Stefanovic, N., et al. (2000) "Object-Based Selective Materialization for Efficient Implementation of Spatial Data Cubes". In *IEEE TKDE*, v.12, n.6, p.938-958.

Stockinger, K. and Wu, K. (2007) "Bitmap Indices for Data Warehouses". In: Data Warehouses and OLAP: Concepts, Architectures and Solutions, Edited by Robert Wrembel and Christian Koncilia. IRM Press, p.157-178.

Wu, K., Otoo, E. J. and Shoshani, A. (2006) "Optimizing Bitmap Indices with Efficient Compression". In *ACM TODS* v.31, p.1-38.

Wu, K., Stockinger, K. and Shoshani, A. (2008) "Breaking the Curse of Cardinality on Bitmap Indexes". http://crd.lbl.gov/~kewu/ps/LBNL-173E.pdf, July.

On Creating a Spatial Integration Schema for Global, Context-aware Applications[1]

Steffen Volz, Daniela Nicklas[2], Matthias Grossmann[3], Matthias Wieland[4]

[2]Carl von Ossietzky Universität Oldenburg, Department of Computing Science, Escherweg 2, 26121 Oldenburg, Germany

[3]Universität Stuttgart, Department of Applications of Parallel and Distributed Systems, Universitätsstr. 38, 70569 Stuttgart, Germany

[4]Universität Stuttgart, Institute of Architecture of Application Systems, Universitätsstr. 38, 70569 Stuttgart, Germany

steffenvolz@gmx.de, daniela.nicklas@uni-oldenburg.de, grossmann@ipvs.uni-stuttgart.de, wieland@iaas.uni-stuttgart.de

Abstract. *The world of spatial data is split into individual data source islands that have different thematic or spatial focuses. When attempting to integrate those data sources, severe challenges arise, since for most GIS application domains a spatial integration schema does not exist. This is also true for the newly emerging domain of mobile, context-aware applications. Since the users of these systems are mobile, transborder access to spatial data or context models is crucial for global deployment. The basis for this work is the Nexus Augmented World Schema, a conceptual schema that serves as an integration standard for autonomous spatial context servers. This paper analyzes some major spatial data standards, especially with respect to the requirements of a spatial integration schema for context-aware applications and illustrates the Nexus approach.*

1. Introduction

The world of spatial data is still split into individual database islands that have different thematic focuses, like e.g., topographic, cadastral, environmental, or traffic related databases. Some are only available for certain geographical regions, e.g., for a city, a federal state, or a country. Handling spatial data across administrative boundaries is rarely possible without difficulties. Hence, a joint use of the separate data islands is often not feasible. This is a big disadvantage regarding the use of GIS in today's decision processes. One of the main integration problems is the lack of a common data schema—which we call spatial integration schema here—that can provide a basis for data integration.

On the other side, a new application domain has emerged in research and industry: So-called context-aware (and often location-based) applications adapt their behavior depending on the situation of their user or the physical world (Schilit, Adams & Want 1994), e.g., navigation systems or ubiquitous computing environments envisioned by Weiser (1991). Common to these applications is their need for context information that

[1] This work was funded by the Collaborative Research Center *Nexus: Spatial World Models for Mobile Context-Aware Applications* (grant SFB 627).

can be used to derive the user's current situation. Depending on the application, such a context model contains various kinds of information. An analysis of different applications shows that their needs for data are overlapping. Common data types in context models are geographical context like digital map data, technical context like available devices or communication networks, information context like points of interest or documents relevant to the current place or situation, and dynamic context from sensors like mobile objects or weather data. By maintaining common context models and sharing them between applications we can significantly reduce the development and deployment effort of context-aware applications.

This paper is organized as follows: In Section 2, we illustrate the vision of a global, federated context model that can be achieved with a spatial integration schema. Section 3 covers related work. In Section 4, we analyze existing spatial data standards with respect to the criteria that were defined in Section 2.1. Finally, in Section 5, the Nexus Augmented World Schema (AWS), a conceptual spatial schema that solves the proposed problems is described.

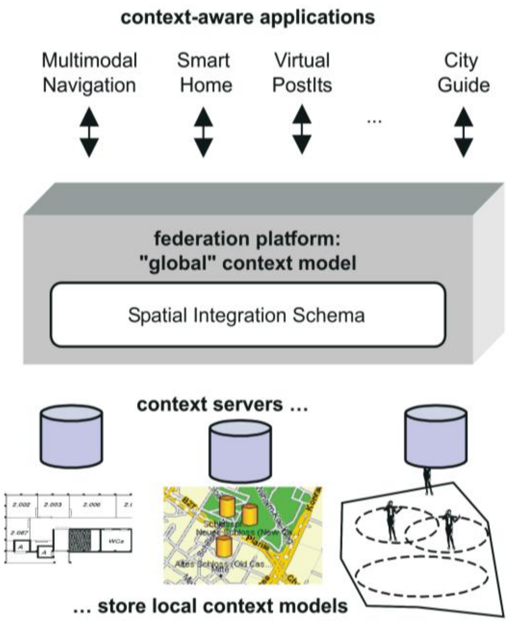


Figure 1. A global, federated context model

2. The Vision of a Global and Federated Context Model

The goal of the Nexus project is to support all kinds of context-aware applications through an open platform that efficiently manages large-scale context models. The platform is realized as a federation between so-called context servers that store local context models. This allows autonomous data providers keeping control of their context data. Using this platform we increase both the total amount of available information and the geographical scope. Figure 1 depicts the high-level architecture of a global context model. We assume that each context server has only data for a certain geographical region and of certain object types. Most of the data objects contain some spatial features.

To achieve this goal, we develop a federation platform that provides applications a single interface and an integrated view on the context data. One important problem that has to be solved is to find an integration schema that facilitates this task. The next section states important requirements or criteria that such an integration schema has to fulfill.

2.1. Criteria

The criteria listed below have been selected as the most important indicators that allow an answer to the question of the suitability of an existing standard for the requirements of global and federated context models:

Support for multiple representations: In a system integrating a large number of independent data providers, it is unavoidable that their data models overlap both by geographical scope and content. The result of a query, e.g., a selection of buildings within a given area may contain different representations of the same real-world object. The data schema should be able to represent this fact, and the system should provide a way to handle multiple representations, e.g., by merging them to a single object when possible.

Globally unique IDs and locators: Objects need some kind of unique ID in order to be referenced by applications. If a system supports merging different data sets from arbitrary sources, IDs for new objects have to be generated in a way that their global uniqueness is assured. Additionally, a federated system should provide some way for an application knowing only the ID of an object to retrieve the object itself, i.e., to determine the object's data provider.

Support for dynamic data: Most location-based applications do not only need static data like roads or hotels, but also dynamic data of moving objects like people or vehicles. Additionally, applications may require dynamic data generated by sensors. Thus, such objects must be modeled appropriately within the data schema and the system must be able to support these objects.

Defined schema semantics: To do anything useful with the data received, applications must know the semantics of the data, i.e., an application receiving an object of type 'building' has to understand what the data provider means with 'building'. The most simple (and most efficient) way to achieve this is to use defined schema semantics.

Schema extensibility: It is obviously impossible to define a global schema that fulfills all requirements of present and future data providers. Thus, the system has to provide a means to extend the schema dynamically. Existing applications, which do not support the schema extension, should nevertheless be able to make use of the data belonging to the schema extension in some way.

Conformity to OGC/ISO-TC211: The Open Geospatial Consortium (OGC) and the ISO-TC211—which are meanwhile closely coupled—published standards for geodata schemas and exchange languages. While they may not be suitable for all applications, embedding parts of them in other systems where feasible may increase interoperability of those systems.

Internationality: Some of today's geodata schemas are specifically tailored to administrative structures of the country they were originally designed for. This is a major disadvantage for a system of global extent, as converting spatial data from other countries is very difficult and usually leads to some loss of information. Ideally, the system would provide a basic schema, which reflects the most important features for all countries and which allows some kind of extensibility to be able to represent more specific features.

Supported types of application: Some geodata schemas were designed with a specific application in mind. Using those schemas in different application domains is usually difficult or even impossible. Again, an ideal schema would provide the capability to model the basic aspects relevant to most applications while being extensible for more specific requirements.

Representing history and prognosis: Applications like traffic jam prognosis or a city guide integrating historic data require the data schema to be able to simultaneously represent the different states of an object at different times in order to deduce the evolution of an object. Additionally, also propositions of an objects' future behavior must be possible.

3. Related Work

Information integration systems have been a research topic for many years. Today, commercial enterprise application integration (EAI) products like IBM's DB2 Information Integrator (Bruni 2003) as well as some research prototypes like TSIMMIS (Garcia-Molina et al. 1997) focus on integrating a previously known fixed set of data sources and on developing an integrated schema for these sources. The area of schema integration (Batini, Lenzerini & Navathe 1986) as well as software systems for this purpose, e.g., Clio (Miller et al. 2001), have a similarly restricted focus as they address the problem of transforming a fixed set of source schemas (and associated data) into an optimally calculated or predefined target schema. They lack the flexibility to handle dynamically changing sets of data providers and they do not aim at semantically integrating the data of different providers. EAI and schema integration software can be used locally to transform the data of a provider into a common domain schema compliant representation. Additionally, none of these approaches takes the special characteristics of spatial data into account.

In the field of GIS, approaches to overcome semantic heterogeneities between different application domains or user communities, respectively, have been made using ontologies or semantic mappers (translators). For example, Fonseca, Egenhofer, Agouris & Câmara (2002) presented an architecture for Ontology-Driven GIS (ODGIS) where objects can be extracted from geospatial databases according to ontological criteria like roles, functions or constituting parts which are stored for each object type. Bishr, Pundt & Rüther (1999) proposed a method that uses domain-specific interfaces in order to map spatial data from one community's ontology (e.g., topography-centered) to the ontology of another one (e.g., transportation-centered). Cruz, Rajendran, Sunna & Wiegand (2002) also used an ontology-based approach in their system. Queries are posed against a global schema (the ontology) and are rewritten by the query processor with the aid of expert-provided agreement files, which describe the mapping between the global and the local schema. A specific problem of information integration – the problem of multiple representation databases – was discussed by Balley, Parent & Spaccapietra (2004). The authors present the situation in France where they have three coexisting spatial databases containing the same real world objects. It is shown how these multiple representations can be modeled in a global geospatial data schema that allows global querying and ensures global consistency. None of the approaches in the

GIS domain has up to now been dealing with integration issues for global, context-aware applications. The modeling of different types of context and the federation of context models have not been tackled as well.

The GIS mediation system proposed by Boucelma, Essid & Lacroix (2002) integrates data coming from a set of Web Feature Servers (WFS). The WFSs can use different local schemas; the mediation system uses a global-as-view approach to integrate the schemas. While this approach simplifies rewriting the queries for the WFS by the mediation systems, it partly exposes the schema heterogeneity to the applications. Adding new WFSs usually require changes of the global schema, which limits the scalability of this approach with respect to the number of data sources.

Finally, a number of integration frameworks for context-aware applications have been developed over the last years. Barretto & da Silva (2004) give a good overview of these approaches and proposes an own approach. This work focuses on system integration in ubiquitous computing. In contrast, we focus in this contribution on the data modeling and data integration, particularly for spatial data models.

4. Existing Spatial Data Standards

Spatial data standards have been defined by different institutions all over the world either from different application perspectives or as a general fundament for different kinds of applications. Basically, they can be divided into international standards specified by global standardization organizations and national standards defined by individual countries for their specific purposes. Both types of standards shall be considered in the following sections.

In the last decade, efforts have been made by the Open Geospatial Consortium (OGC) and the International Standards Organization (ISO) in close collaboration to achieve interoperable GIS on an international level. Up to now, mainly technical interoperability issues were taken into account and as one of the main results the Geography Markup Language (GML) was developed (Portele 2007). However, semantic interoperability is a topic that has not been addressed sufficiently by standardization organizations although its importance has been recognized. The OGC has also started an initiative called OpenLS (Open Location Services) that mainly specifies services useful for location-based applications but does not provide a detailed model of the world.

Another spatial data standard of international relevance which was originally invented by the industry and meanwhile has become an ISO standard as well is the Geographic Data Files Format (GDF) (ISO 14825:2004). It was developed for the purpose of car navigation systems.

There are also country-specific standards which are driven by national organizations like the Federal Geographic Data Committee (FDGC) in the United States or the Spatial Information Council for New Zealand and Australia (ANZLIC). They are aiming at the development of geospatial data infrastructures. In Germany, the Working Committee of the Surveying Authorities of the States of the Federal Republic of Germany (AdV) has established national standards, namely ATKIS (Authoritative Topographic Cartographic Information System) and ALK (Automatic real estate map). ATKIS and ALK are being transformed into a universal data schema for all kinds of authoritative surveying data of Germany, called the AFIS-ALKIS-ATKIS reference model (AFIS: reference point information system, ALKIS: Authoritative real estate information system) by which the object definitions within the separate schemas of ALK an ATKIS are harmonized (Working Committee of the Surveying Authorities of the States of the Federal Republic of Germany 2008). Another standard called CityGML is just evolving in Germany. It has been initiated by the Special Interest Group 3D of the spatial data infrastructure project of North Rhine-Westphalia and defines an application schema based on GML to represent 3D city models at different levels of detail (Kolbe, Gröger & Plümer 2005).

4.1. Evaluation of Existing Standards

In this section we discuss to which degree the available standards meet the needs of a spatial integration schema for context-aware applications. This kind of evaluation is sometimes ambiguous and difficult to do. Thus we set up 4 categories to express tendencies for the suitability of a standard for the respective task/requirement: Impossible/ not supported: – ; partially (but not entirely) possible/supported: (–); possible/supported using work-arounds: (+); possible/supported: +.

The results of the evaluation are presented in Table 1. The last column of Table 1 contains the evaluation of the Augmented World Schema (AWS) which is discussed in Section 5.

Multiple representations / geometries. Typically, no existing standard can express that two data objects in different data sets represent the same real world entity. Hence, multiple representations cannot be described by existing standards at all, and they always exist separately, without knowledge of each other. However, at least in ATKIS and AAA, an object of the real world can be represented by different geometries, e.g., a place can either be defined as a point or as an area. Also in CityGML, due to its level of detail approach, an object can be represented by different geometries, e.g., only as a cuboid or as a detailed 3D object with roofs etc. On the other hand, GDF data are only captured in approximately the same scale and its objects cannot contain multiple geometries.

Globally unique ID. AAA, ATKIS and GDF are aiming at system-wide unique IDs whereas ALK and CityGML do not provide mechanisms supporting this feature. None of the schemas offers locators, i.e., a concept by which an object can be found within a distributed system when the identification of the object is known.

Support for dynamic data. GDF, as well as all the other presented standards, does not contain structures for dynamic or mobile objects and temporal relations, i.e., cars, pedestrians, etc. and chronologies expressing what happened before, during or after a certain event cannot be represented in schemas of existing standards.

Defined schema semantics. Except ALK, which does not have a strong semantic description behind it, all of the presented standards do have a fixed schema thoroughly defining the properties of the comprised objects.

Schema extensibility: Currently, only CityGML supports application specific extensions due to its GML origin. All other spatial data schemas do not foresee to extend the existing classes. E.g., introducing mobile objects or footpaths within GDF, ALK or ATKIS would result in non-standard specific extensions which could not be used by other applications. Instead, such changes of the original data schema would only be realizable by tedious standardization processes.

Table 1. Evaluation of existing spatial data standards

Requirements/Standards	ATKIS	ALK	GDF	AAA	CityGML	AWS
Multiple representations	–	–	–	–	–	(+)
Multiple geometries	(+)	–	–	(+)	+	+
Globally unique IDs	(+)	–	(+)	(+)	(–)	+
Support for dynamic data	–	–	(–)	(–)	–	+
Defined schema semantics	+	(–)	+	+	+	+
Schema extensibility	–	–	–	–	(+)	+
Supported types of applications	(+)	–	(–)	(+)	(–)	(+)
Conformity to OGC/ISO-TC211	–	–	–	(+)	+	(–)
Internationality	–	–	(+)	(–)	(+)	(+)
History and prognosis	–	–	(–)	(+)	(–)	(+)

Supported types of application. Adapting existing standards to multiple applications is hardly achievable. With respect to the requirements of global context models for context-aware applications (e.g., for performing intermodal navigation where a user can traverse multiple networks like roads as well as public transportation lines and pedestrian paths), GDF is too much restricted to road data and does—at the current stage—not contain essential object classes, at least not in necessary detail. The same is true for ALK, ATKIS, AAA and CityGML. Although ATKIS and AAA show more comprehensive conceptual schemas that were developed as a basic schema for the representation of topographic objects being able to support diverse applications, they are still lacking some object classes necessary for context-aware services like sensors. Furthermore, for one important context aware application—navigation—ATKIS and AAA data are not suitable: Information about how roads can be traversed or about driving limitations (forbidden maneuvers) is missing.

Conformity to OGC/ISO-TC211. The central standardization organizations for spatial data are OGC and ISO-TC211. Up to now, semantic interoperability has not been tackled sufficiently by OGC/ISO, but at least important technical obstacles have been overcome that should be adopted by the GIS community. Since ATKIS and ALK are standards which were developed at the beginning of the 80s, they of course do not consider international standardization issues. However, the AAA model uses only OGC or ISO definitions to represent the geometries of its objects and GML to exchange the data. GDF, although heading towards X-GDF where things are planned to be changed (e.g., it is intended to extend the schema to be able to perform pedestrian navigation and to represent 3D city models or to adjust the schema to OGC/ISO standards) (van Essen & Hiestermann 2005), currently does not adhere to OGC/ISO-TC211 guidelines.

Internationality. Since ALK, ATKIS and AAA are restricted to Germany, they can only be regarded as national standards without global relevance. In contrary, GDF includes internationalization issues, e.g., different address formats for different countries. CityGML also defines rather general object structures which can be used all over the world for the modeling of urban areas.

History and prognosis. Different temporal states of objects can only be expressed appropriately by the AAA reference model as it proposes a versioning concept which allows storing different versions of an object within an object container and so object histories can be reconstructed. GDF allows time-dependent attributes of objects. CityGML does not support different temporal representations besides those that can already be handled by GML. ATKIS and ALK are completely neglecting this issue. None of the schemas presents features to describe future states of an object derived by simulation or prognosis algorithms.

As a result of the evaluation, we found that no existing spatial data standard can be used to realize global and federated context models without facing significant drawbacks. Hence, we set up an integration scheme that optimally supports the domain of context-aware applications in such an environment: the Augmented World Schema.

5. Our Solution – The Augmented World Schema

The Augmented World Schema (AWS) is an object-oriented, common data schema used by context servers, the Nexus federation and context-aware applications to exchange and process data (Nicklas & Mitschang 2004). A fixed set of classes (the Standard Class Schema (SCS)) comprises those classes we consider basic for most context aware applications. Figure 2 shows an excerpt of the SCS on the left.

Object classes of the AWS are modeled in a class hierarchy. All spatial objects inherit from the basic object class SpatialObject which ensures that every AWS object has a unique identifier (the NOL, see below) and a geographic position. Additionally, every AWS object has a special attribute called type, which contains the name of the object class. With that, the object class can be queried like any other attribute of the object. The AWS offers some flexibility that extends the standard object-oriented approach: Many attributes are marked as optional, i.e., providers may choose to supply those attributes or not, but if they are supplied, their name and semantics are defined. The middleware expects that every application is able to handle SCS objects.

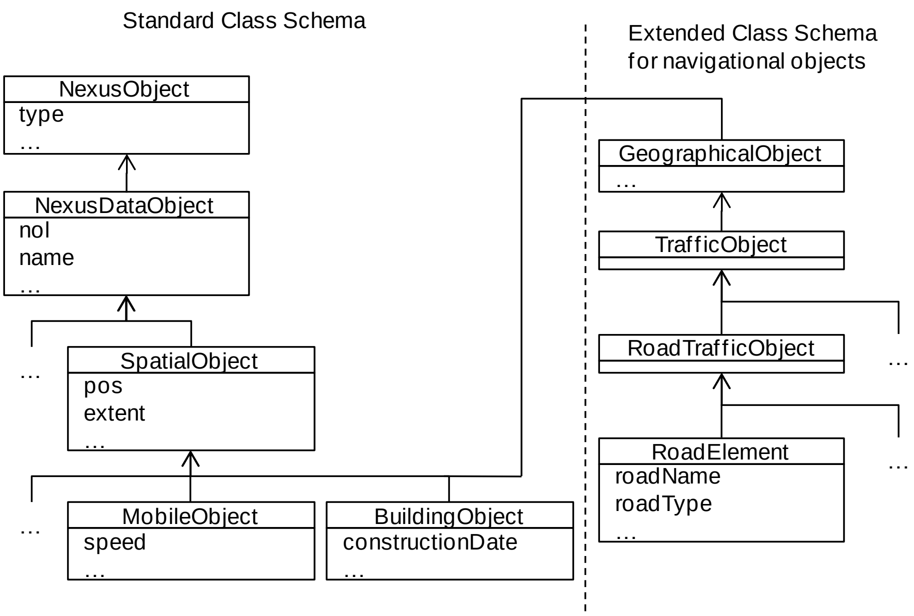


Figure 2. Standard Class Schema and an Extended Class Schema

The data of the AWS (i.e., the Augmented World Model) is not materialized at the federation tier but is distributed across the data providers. For this, we use the concept of Augmented Areas (AA) which corresponds to a certain geographical region and is associated with the SCS and optionally with one or more extended class schemas. The regions of different AAs may overlap each other.

5.1. Realization of Requirements

In the following, it is depicted how the criteria that have been set up in Section 2.1 as the basic requirements for a spatial integration schema are realized by the AWS:

Multiple representations. Since Augmented Areas can be overlapping both in space and in content, the AWS has to allow multi-represented objects. By means of explicit relations called MultiRepresentational Relations (or MRep Relations in short), multiple representations can be linked and made aware of each other (Volz & Walter 2004). As MRep Relations refer to a spatial location that can be derived from the geometries of the related objects, they can be spatially indexed and stored on regular, spatially-enabled context servers within Nexus. The concept of MRep Relations enables integration processes carried out by the federation component. This shall be illustrated by means of an example: If a client query requests street data for a certain geographical region and this request cannot be satisfied by one single Augmented Area alone, integration mechanisms are triggered on the federation level. This involves two steps: First, the street data of the two Augmented Areas plus the MRep Relations for those objects in the overlapping area which are multiply represented have to be loaded from the appropriate context servers. Secondly, the federation can merge the multiple representations into combined objects by exploiting MRep Relations (Volz & Bofinger, 2002). First approaches have been implemented here which geometrically adjust the data based on rubber-sheeting techniques and produce an average geometry from the two source geometries based on a triangulation. The result of a merge operation can be an object that has multiple types and multiple instances of the same attribute.

Globally unique ID. Each object in the Augmented World model has an attribute called Nexus Object Locator (NOL), which can be used both for identifying an object and finding the server where the object is stored. The NOL consists of three main parts: the object ID, the Augmented Area ID (both are globally unique) and a service description containing information on the service hosting the object (e.g., a web service). If two objects have the same object ID, they are considered to be representations of the same real-world entity (like having a MRep relation). This can be used to deliberately distribute different aspects of one object to different context servers, e.g., to optimize the management of the dynamic and the static parts of a mobile object (Grossmann, Bauer, Hönle, Käppeler & Nicklas 2005). Also, a data provider that wants to offer additional information to a given data set of another provider can use the object IDs of that provider for its own data objects (e.g., to add sightseeing information to given map data).

Support for dynamic data. The AWS contains static, dynamic, and mobile data objects. For the modeling, there is no difference between static and dynamic data: Both can be modeled as attribute values of AWS objects. However, a meta data concept (Hönle et al. 2005) allows representing additional information for a certain datum. This is especially designed for handling dynamic data, where meta data like time stamps, accuracy, update rate etc. are needed by applications. Mobile objects often change the data value of their spatial attributes (i.e., position), which is one of main selection criteria for data objects in that application domain. This afflicts both the data management and the applications. Therefore, the AWS models mobile objects in an own hierarchy (i.e., all mobile objects inherit from the class 'MobileObject').

Common schema semantics. This aspect is covered by class schemas. The identification of required object classes for the AWS has been based on a use case analysis for context-aware applications on the one hand and on the other hand on an analysis of existing (external) conceptual schemas. For these external schemas, we developed matching techniques in order to identify semantically corresponding object classes amongst them (Volz 2005) and applied mechanisms to integrate these object classes into the global schema. For example, in the case of existing conceptual schemas for street data, we established very basic object classes in the AWS like JunctionElement, RoadElement, DirectedRoad, etc. and neglected the complex data structures available in existing schemas for street data. Thus, we were able to develop mapping functions in order to transfer street data from existing databases into object classes provided by the AWS, i.e., the context servers of Nexus can easily be enriched by existing geospatial data applying these mapping functions.

Schema extensibility. By means of Extended Class Schemas, additional types can be introduced in the AWS, as shown in Figure 2 (on the right). If multiple applications share this knowledge, they have again a common schema semantic. The SCS may be extended by so called Extended Class Schemas (ECS). The object types of the ECS can inherit from any type in the SCS. ECSs may be defined anytime by any provider, so they are not necessarily known to all applications or to the middleware. We address this problem by two different approaches: First, applications specify with each query the ECS they know about. If the result of the query contains an object whose type belongs to a different ECS, it gets cast into the closest class of the SCS. And secondly, the middleware is able to extract information from structures it does not know completely.

Supported types of applications. The AWS is designed to support all kinds of context-aware applications, with an emphasis on location-awareness. However, most of the classes define objects of the real world with rather common attributes. Hence, data from the AWS can be used for other application domains needing spatial data.

Conformity to OGC/ISO-TC211. For several reasons, AWS is no GML Application Schema—we found that some of the concepts of a feature did not match the requirements of the AWS, especially with regard to multiple representations. However, for representing spatial and temporal data types, we use GML schemas.

Internationality. Since Nexus is designed as a global spatial information platform, we strived for an international design. However, we are aware that different cultures do have different views on the world. Therefore, the modeling of the AWS fits best in the western world.

History and prognosis. By means of the already mentioned meta data, timestamps and other temporal information can be added to the data objects and attribute values. The query language also provides temporal predicates to use this information to select and restrict result sets which already allows querying past states of the Augmented World Model. In the future, we plan to use data warehouse and data mining techniques to exploit this information for prognoses, e.g., in traffic analyses.

5.2. System Support

For supporting context-aware applications, developing a data schema only is not sufficient, a system for maintaining the data is also necessary. In large-scale scenarios, simple approaches like storing all the data on a single web server are not feasible. Instead, a distributed infrastructure is necessary, which on the one hand allows disseminating the data across many different servers and on the other hand provides a homogenous, integrated view of the data to the applications. The Nexus system (Nicklas et al. 2001) provides a proof-of-concept implementation fulfilling the requirements proposed in Section 2.1 using a three-tier architecture.

6. Conclusion

In order to realize the federation of context models in distributed environments with heterogeneous sources, a spatial integration schema has to exist. In this paper, we defined the basic requirements that such a schema has to conform to. Since context models are mainly organized in a spatial manner, we investigated existing spatial data standards and matched them against the defined requirements. We found out that—although we analyzed different types of standards—none of them was optimal since they were all designed from different viewpoints and were targeted to accomplish other tasks. As a consequence, we set up an own schema definition which conforms to existing standards as far as possible but on the other hand provides necessary extensions to support federated context models.

References

Balley, S., Parent, C. & Spaccapietra, S. (2004). Modelling geographic data with multiple representations. *Intl. Journal of Geographic Information Systems 18*(4), *pp. 327-352*

Batini, C., Lenzerini, M. & Navathe, S.B. (1986). A Comparative Analysis of Methodologies for Database Schema Integration. *ACM Computing Surveys, 18*(4)

Barretto, S., & da Silva, M.M. (2004). Using Integration Frameworks for Developing Context-Aware Applications. *Proc. of the 2nd European Symposium on Ambient Intelligence (EUSAI), LNCS 3295, Springer-Verlag, pp. 100-111*

Bishr, Y.A., Pundt, H. & Rüther, C. (1999). Proceeding on the road of semantic interoperability—Design of a semantic mapper based on a case study from transportation. *Proc. of the 2nd Intl. Conference on Interoperating Geographic Information Systems, Zurich, LNCS 1580, Springer-Verlag, pp. 203-215*

Boucelma, O., Essid, M. & Lacroix, Z. (2002). A WFS-based mediation system for GIS interoperability. *Proc. of the* 10^{th} *ACM Intl. Symposium on Advances in Geographic Information Systems, ACM, pp. 23-28*

Bruni, P. (2003). Data federation with IBM DB2 Information Integrator V8.1. *IBM, International Technical Support Organization*

Cruz, I., Rajendran, A., Sunna, W. & Wiegand, N. (2002). Handling semantic heterogeneities using declarative agreements. *Proc. of the* 10^{th} *ACM Intl. Symposium on Advances in Geographic Information Systems, ACM, pp. 168-174*

Fonseca, F., Egenhofer, M., Agouris, P. & Câmara, G. (2002). Using ontologies for integrated geographic information systems. *Transactions in GIS 6*(3)*, pp. 231-257*

Garcia-Molina, H., Papakonstantinou, Y., Quass, D., Rajaraman, A., Sagiv, Y., Ullman, J. D., et al. (1997). The TSIMMIS approach to mediation: data models and languages. *Journal of Intelligent Information Systems 8*(2)

Grossmann, M., Bauer, M., Hönle, N., Käppeler, U., Nicklas, D. & Schwarz, T. (2005). Efficiently managing context Information for large-scale scenarios. *Proc. of the 3rd IEEE Intl. Conference on Pervasive Computing and Communications, IEEE Computer Society, pp. 331-340*

Hönle, N., Käppeler, U., Nicklas, D., Schwarz T. & Großmann, M. (2005). Benefits of integrating meta data into a context model. *Proc. of 2nd IEEE PerCom Workshop on Context Modeling and Reasoning, IEEE Computer Society*

Kolbe, T. H., Gröger, G. & Plümer, L. (2005). CityGML—Interoperable access to 3D city models. *Geo-Information for Disaster Management (GI4DM), Springer-Verlag*

Miller, R.J., Hernández, M.A., Haas, L.M., Yan, L., Ho, C.T.H., Fagin, R. & Pope, L. (2001). The Clio Project: Managing Heterogeneity. *ACM Press, SIGMOD Rec. 30*(1)

Nicklas, D., Großmann, M., Schwarz, T. Volz, S. & Mitschang, B. (2001). A Model-Based, Open Architecture for Mobile, Spatially Aware Applications. 7th Intl. Symposium on Advances in Spatial and Temporal Databases (SSTD)

Nicklas, D. & Mitschang, B. (2004). On building location aware applications using an open platform based on the Nexus Augmented World Model. *Software and Systems Modeling*, *3*(4)*, pp. 303-313*

Portele, C. (2007). OpenGIS Geography Markup Language (GML) Encoding Standard, Version 3.2.1. *http://portal.opengeospatial.org/files/?artifact_id=20509*

Schilit, B.N., Adams, N. & Want, R. (1994). Context-aware computing applications. *IEEE Workshop on Mobile Computing Systems and Applications, IEEE Computer Society*

van Essen, R. & Hiestermann, V. (2005). "X-GDF" — The ISO Model of Geographic Information for ITS. *ISPRS Workshop on Service and Application of Spatial Data Infrastructure, XXXVI(4/W6)*

Volz, S. & Bofinger, J. (2002). Integration of spatial data within a generic platform for location-based applications. *Proc. of the Joint Intl. Symposium on Geospatial Theory, Processing and Applications*

Volz, S. & Walter, V. (2004). Linking different geospatial databases by explicit relations. *Proc. of the XXth ISPRS Congress, Comm. IV, pp. 152-157*

Volz, S. (2005). Data-driven matching of geospatial schemas. *Proc. of the 8th Conference on Spatial Information Theory (COSIT), LNCS 3693, Springer-Verlag, pp. 115-132*

Weiser, M. (1991). The computer for the 21st century. *Scientific American, Sep. 1991*

Working Committee of the Surveying Authorities of the States of the Federal Republic of Germany (2008). Documentation on the modelling of geoinformation of official surveying and mapping in Germany (GeoInfoDok), Version 6.0

Metodologia para Representação da Estrutura de Dados Geoespacial Vetorial da Mapoteca Nacional Digital em Bancos de Dados Geográficos Relacionais

Rogério Luís Ribeiro Borba[3], Guilherme L. A. Mota[1,3], Jorge Luís Nunes e Silva Brito[2,3]

[1]Departamento de Engenharia de Sistemas e Computação
[2]Departamento de Engenharia Cartográfica
[3] Programa de Pós-graduação em Engenharia de Computação - Geomática
Universidade do Estado do Rio de Janeiro
Rio de Janeiro – RJ – Brasil

rogerioluisribeiro@yahoo.com,{guimota,jsilvabr}@gmail.com

Abstract. *A single model of vector geospatial data structure that could be used by institutions responsible for systematic land mapping, has always been a major challenge in mapping national scene. The National Commission of Cartography, in 2007, launched the technical specifications for geospatial data vector that finally unite the form of mapping models and break with old paradigms. One of these specifications is a conceptual object oriented model. This paper presents a solution to map this conceptual model for a generic model of relational geographic database.*

Resumo. *Um modelo único de estrutura de dados geoespacial vetorial que pudesse ser utilizado pelas instituições responsáveis pelo mapeamento sistemático terrestre, sempre foi um dos grandes desafio no cenário cartográfico nacional. A Comissão Nacional de Cartografia, em 2007, lançou as especificações técnicas de dados geoespaciais vetoriais que finalmente unificam a forma de fazer mapeamento e quebram paradigmas com modelos antigos. Uma destas especificações é um modelo conceitual orientado a objeto. O presente trabalho descreve uma solução para mapear este modelo conceitual para um modelo genérico de banco de dados relacional geográfico.*

1. Introdução

No Brasil há duas instituições públicas responsáveis pelo mapeamento topográfico sistemático terrestre. O Exército Brasileiro, através da Diretoria de Serviço Geográfico (DSG), e o Instituto Brasileiro de Geografia e Estatística (IBGE). Até o início dos anos 90, a chamada cartografia convencional ou analógica era executada, pelas duas organizações, de forma padronizada, segundo especificações técnicas obrigatórias, de acordo com resoluções definidas pelos órgãos responsáveis, em obediência a legislação específica (Delou, 2006).

A partir da década de 90, a tecnologia digital passou a ser utilizada pelos órgãos encarregados da produção cartográfica. O crescimento da produção de bases cartográficas digitais por estes órgãos, sem normatização, causou a ruptura do processo padronizado de produção cartográfica. Cada instituição criou sua própria padronização para a estruturação dos elementos cartográficos contidos nos arquivos vetoriais por ela gerados. Assim, apesar das facilidades e avanços proporcionados pela introdução da tecnologia digital, passa a haver dissonância no processo de produção cartográfica digital. Enquanto a DSG baseava a estruturação dos elementos cartográficos nas Tabelas da Base Cartográfica Digital – TBCD; o IBGE baseava a estruturação dos elementos cartográficos na Mapoteca Topográfica Digital – MTD. Em função desta separação de modelos, diversos problemas foram gerados, conforme observado por Delou (2006).

No ano de 2004, a Comissão Nacional de Cartografia (CONCAR), por intermédio da Subcomissão de Estruturação de Dados Espaciais (SDE) e do Comitê de Estruturação da Mapoteca Nacional Digital, deu início à definição de um novo modelo para unificar a estruturação de elementos cartográficos para a cartografia no Brasil, o que aconteceu em 2007, tendo sido lançadas as "especificações técnicas para estruturação de dados geoespaciais vetoriais (EDGV)" [CONCAR 2007]. A EDGV é parte componente da Mapoteca Nacional Digital (MND). A MND é componente da estruturação de dados cartográficos do Mapeamento Sistemático Terrestre, da Infra-estrutura Nacional de Dados Espaciais (INDE) [CONCAR 2007].

A EDGV foi modelada segundo o paradigma da orientação a objetos (OO), uma vez que esta abordagem é a mais adequada para criação de um modelo do mundo real. Além disto o uso da abordagem OO foi consenso entre os membros da CONCAR, dentre os quais se destacam a DSG e o IBGE.

A princípio, tanto as aplicações quanto as bases de dados deveriam seguir o paradigma da OO. Contudo, no que se refere ao armazenamento do dados, o modelo relacional é adotado na grande maioria das empresas devido à sua maturidade e seu formalismo matemático. Além disso, os investimentos realizados para a aquisição de sistemas gerenciadores de bancos de dados relacionais (SGBDR) e treinamento de pessoal teriam que ser refeitos no caso da migração para bancos de dados OO. Pelas razões anteriormente listadas, faz-se necessária a definição de uma estratégia para a implementação da EDGV em SGBDR e, assim, possibilitar a utilização da infra-estrutura atualmente instalada. É preciso ainda destacar dois requisitos adicionais desta demanda. Primeiro, o fato deste mapeamento ser feito em bases de dados geoespaciais, portanto, considerando as componentes geométricas. Segundo, a estratégia de mapeamento deve ser genérica de forma a permitir sua implementação em quaisquer SGBDR que possuam extensão espacial.

A presente iniciativa faz parte do projeto E-FOTO que desenvolve uma estação fotogramétrica digital educacional livre (licença GNU/GPL). A intenção é criar condições para que a restituição fotogramétrica digital do E-FOTO possa ser armazenada em SGBDR livres ou proprietários.

A seguir, na seção 2, é apresentada brevemente a EDGV. Na seção 3, é abordada a componente geométrica; na seção 4, é apresentada a metodologia proposta e o mapeamento em SGBDR; na seção 5, são apresentados os resultados obtidos. Por fim, a seção 6 apresenta as conclusões da pesquisa.

2. Estruturação dos dados geoespaciais vetoriais

As especificações técnicas da EDGV definem um padrão único a ser seguido pelas instituições responsáveis pelo mapeamento topográfico sistemático terrestre. No contexto da presente proposta, dentre as especificações da EDGV, o modelo conceitual OO é a mais importante. Esse modelo é organizado em treze categorias (pacotes-conforme UML): (1) Abastecimento de Água e Saneamento Básico; (2) Administração

Pública; (3) Educação e Cultura; (4) Energia e Comunicações; (5) Estrutura Econômica; (6) Hidrografia; (7) Limites; (8) Localidades; (9) Pontos de Referência; (10) Relevo; (11) Saúde e Serviço Social; (12) Sistema de Transportes; (13) Vegetação.

Cada categoria/pacote na EDGV contempla um diagrama de classes. Existem basicamente três tipos de classes:

- Classes de Feição - possuem geometria associada (ponto, linha ou polígono);
- Classes Agregadoras - não possuem geometria, somente agregam classes que possuem geometria;
- Classes de Domínio - não possuem geometria e nem agregam outras classes, simplesmente descrevem conjuntos de valores comuns. Podem ser referenciadas por classes do tipo Feição e Agregadora.

O modelo conceitual OO da EDGV após traduzido para o esquema relacional, poderá ser utilizado, por exemplo, no armazenamento de dados geográficos 3D adquiridos através do processo de restituição digital. Portanto, é necessário que o armazenamento de informações alfanuméricas e espaciais seja uniforme. A seção a seguir trata deste suporte por parte dos SGBDR.

3. O suporte a informação geográfica em SGBDR

Muitos SGBDR, como o Postgresql e o Oracle, disponibilizam mecanismos para que, além de informações alfanuméricas, sejam armazenados dados espaciais. Assim, a informação geográfica num registro de uma tabela pode coexistir com os respectivos dados alfanuméricos. Além disto, é importante mencionar que os bancos de dados supracitados, por seguirem os padrões definidos pelo Open Geospatial Consortium (OGC), possibilitam a interoperabilidade no tratamento da informação geográfica.

O modelo geométrico desenvolvido pela OGC tem como classe ancestral a *Geometry* que é abstrata e especializada em outras classes. Assim, é possível tratar adequadamente classes cuja forma geométrica não seja conhecida a priori. Este é o caso da classe Elemento_Fisiografico_Natural da categoria Relevo da EDGV, que pode ser representada como: ponto, linha e polígono. Assim ao mapear esta classe para uma tabela, uma coluna do tipo *Geometry* deve ser criada para tratar sua representação.

Por outro lado, alguns softwares não possuem a capacidade de lidar com mais de uma forma geométrica para uma dada feição. Este é o caso, por exemplo, do ArcGis, onde ao criar um elemento geográfico é preciso definir concretamente qual o tipo esperado de geometria (ponto, linha, polígono entre outros). Nesse caso, não é possível definir um tipo abstrato como por exemplo *Geometry*. Uma solução alternativa seria criar para cada forma geométrica uma tabela e repetir os atributos alfanuméricos.

4. Implementando o mapeamento do modelo conceitual orientado a objeto para o modelo relacional

Uma diferença importante entre os modelos OO e relacional reside nos tipos de relacionamentos possíveis. O modelo relacional possui apenas um tipo de relacionamento, enquanto o modelo OO possui mais de um tipo de relacionamento, cada um com um significado particular. Dentre os relacionamentos da OO, o modelo conceitual da EDGV emprega os mencionados a seguir: (1) associação; (2) agregação; (3) composição; (4) generalização/especialização. Além destes, a EDGV define um novo tipo de relacionamento para tratar a espacialidade. As subseções a seguir tratam do mapeamento em modelo relacional dos relacionamentos citados acima.

4.1. Associação, agregação e composição

São formas de relacionamentos entre classes, parecidas com os relacionamentos entre tabelas no banco de dados relacional, que utilizam os mecanismos de restrição de integridade. No caso mais específico da composição, onde as partes só existem enquanto o todo existir, no modelo OO basta excluir o pai (todo) para que os filhos (partes) também sejam excluídos. No modelo relacional este mecanismo pode ser efetuado colocando na chave estrangeira restrição de integridade como: **on delete cascade** .

4.2. Generalização, Especialização

É um forma de relacionamento natural no modelo OO, porém no modelo relacional não existe tal mecanismo. Para simular o relacionamento de Generalização/Especialização no modelo relacional, há basicamente três estratégias: (1) Tabela por hierarquia; (2) Tabela por classe concreta; (3) Uma tabela por classe.

A presente proposta adotou a estratégia de uma tabela por hierarquia, que prioriza a simplicidade e a performance. Para este caso, todos os atributos são colocados em uma única tabela (Ambler, 2008). Dentre suas vantagens podem ser mencionadas: (1) simples implementação; (2) fácil inclusão de classes; (3) bom suporte a polimorfismo e (4) acesso rápido. Dentre as desvantagens podem ser enumeradas: (1) a modificação de uma classe pode afetar outras classes na hierarquia; (2) desperdício de espaço; (3) pode gerar tabelas muito grandes; (4) colunas não podem ter restrições do tipo *not-null* em alguns casos; (5) necessidade de uma coluna extra, para indicar a classe. Existem alguns casos de herança múltiplo no modelo conceitual da EDGV. Para mapeá-los também pode ser seguida a estratégia de tabela única. A figura 1 abaixo, exemplifica o uso da estratégia uma tabela por hierarquia.

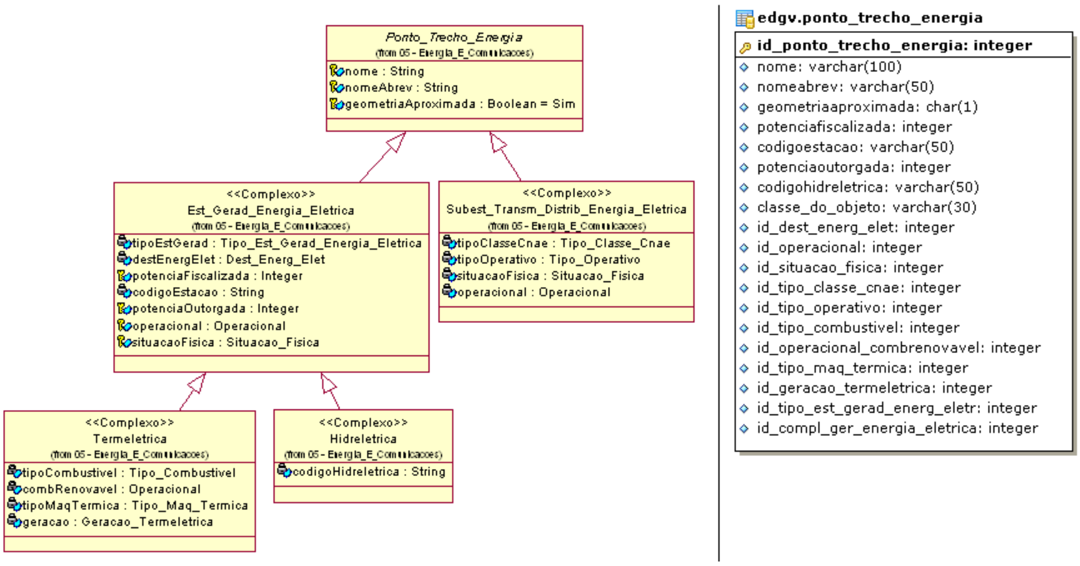


Figura 1. Exemplo da estratégia uma tabela por hierarquia

4.3. Relacionamentos espaciais

No modelo da EDGV além dos relacionamento padrões da orientação a objeto há os relacionamentos espaciais (estereotipo <<espacial>>). Este tipo de relacionamento pode ser implementado diretamente através das operações espaciais dos SGBDR, sem necessitar da definição de relacionamento entre as respectivas tabelas. Um exemplo é apresentado na figura 2.

Neste exemplo, uma área político administrativa contém uma ou mais localidades. Observe que este relacionamento é espacial, ou seja, através de uma operação espacial é possível saber quais localidades pertencem a determinada área político administrativa. Assim, no banco de dados não seria necessário criar relacionamento entre as tabelas.

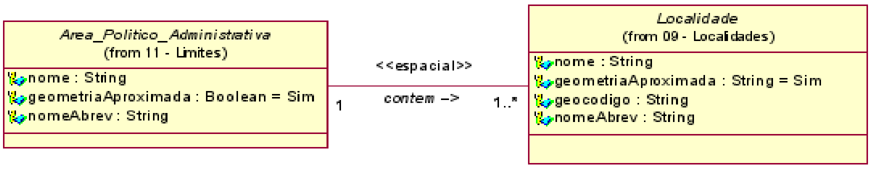


Figura 2 - Relacionamento espacial

5. Experimentos

Os experimentos foram realizados em ambiente de software livre, utilizando o SGBD PostgreSQL com a extensão espacial PostGIS. Portanto, foi necessário o desenvolvimento de scripts SQL. Os seguintes passos foram executados: (1) desenvolvimento e execução de um script para a criação das tabelas de feição; (2) adição de uma coluna, através da função *AddGeometryColumn*, para armazenar a geometria das feições nas tabelas criadas no passo 1 (esta função adiciona informações na tabela de metadados); (3) desenvolvimento e execução de um script para a criação das tabelas derivadas das classes agregadoras e de domínio; (4) desenvolvimento e execução de um script de restrição de integridade para gerar os relacionamentos entre as tabelas. Note que este mesmo procedimento poderia ser adotado para qualquer outro SGBD.

Na figura 3 é mostrada a visualização da inserção de um ponto e um polígono na tabela Elemento_Fisiografico, cuja tipo da coluna geométrica é Geometry, permitindo o armazenamento de um ponto, linha ou polígono.

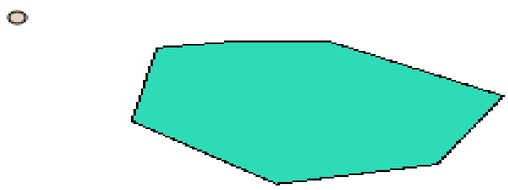


Figura 3. Resultado da inserção de duas feições na tabela Elemento_Fisiográfico visualizado no Software livre QGIS

O resultado alcançado é um esquema relacional/objeto-relacional genérico derivado do modelo conceitual da EDGV, que pode ser utilizado e especializado para PostgreSQL, Oracle Spatial entre outros. A portabilidade deste repositório de dados geoespaciais baseado nos padrões do OGC permite seu compartilhamento com aplicações e softwares geoespaciais.

6. Conclusão

Este trabalho apresentou uma solução para mapear a Estrutura de Dados Geoespacial Vetorial (EDGV), parte componente da MND para bancos de dados geográficos. Neste contexto, os relacionamentos do tipo composição, agregação e associação simples no modelo orientado a objeto da EDGV tornaram-se, no modelo relacional, relacionamentos normais entre tabelas com suas respectivas restrições de integridade. Cada conjunto de classes do relacionamento de generalização/especialização condensou-se em uma tabela com uma coluna especial (classe_do_objeto) para identificar a classe correspondente. Ainda na generalização/especialização os relacionamentos das subclasses foram condensados nesta tabela. Os relacionamentos estereotipados de espacial da EDGV não produziram nenhum efeito no modelo relacional, uma vez que são derivados de funções espaciais. O produto obtido foi um esquema de banco de dados geográfico que será utilizado no software E-FOTO para armazenamento das feições obtida por restituição fotogramétrica digital.

Como conseqüência, este mapeamento disponibiliza um padrão de esquema em banco de dados relacional/objeto-relacional com extensão espacial, permitindo assim uma interoperabilidade em nível de base de dados. Acredita-se que esta solução de geração de esquema possa ser utilizada por qualquer instituição que pretenda realizar seu processos de mapeamento cartográfico em conformidade com a EDGV definida pela Comissão Nacional de Cartografia.

Referências bibliográficas

Ambler, W. Scott. (2008) Mapping Objects to Relational Databases: O/R Mapping In Detail http://www.agiledata.org/essays/mappingObjects.htm [Acessado em agosto].

Booch et. al. (2000)UML Guia do Usuário. Rio de Janeiro: Campus.

Brito, J. L. N. et. al. (2006) Projeto E-FOTO: Disponibilizando uma Ferramenta educacional livre para a Fotogrametria Digital COBRAC 2006 Congresso Brasileiro de Cadastro Técnico Multifinalitário UFSC Florianópolis 15 a 19 de Outubro.

Coelho L. e Brito J. L. N. (2007) Fotogrametria Digital. Rio de Janeiro: UERJ.

CONCAR (2007)Especificações Técnicas para Estruturação de Dados Geoespaciais Digitais Vetoriais. Brasília.

CONCAR (2007)- RESOLUCAO_CONCAR_001-2007. Brasília.

Delou, A. L. (2006.)A. Estruturação de Bases Cartográficas Para SIG: Um estudo comparativo dos modelos do IBGE e da DSG. Rio de Janeiro: UERJ.

Fowler, M. (2002)Analysis Patterns Reusable Object Models. MA.: Addison-Welsey

Lunardi, O. A. & Augusto M. J. C. (2006)Infra-Estrutura dos Dados Espaciais Brasileira – Mapoteca Nacional Digital. COBRAC 2006 · Congresso Brasileiro de Cadastro Técnico Multifinalitário · UFSC Florianópolis · 15 a 19 de Outubro .

Mapoteca Topográfica Digital(1996): Documento Geral – Versão 3.0. IBGE, Diretoria de Geociências, Departamento de Cartografia. Rio de Janeiro.

Proposição de Infra-Estrutura de Dados Espaciais (SDI) Local, Baseada em Arquitetura Orientada por Serviços (SOA)

Pedro A. Oliveira[1], Clodoveu A. Davis Junior[2], Pedro Felipe A. Oliveira[3]

[1]Pontifícia Universidade Católica de Minas Gerais (PUC Minas)
pedroalves@pucminas.br
[2]Universidade Federal de Minas Gerais (UFMG)
clodoveu@dcc.ufmg.br
[3]GeoExplore – Belo Horizonte - MG – Brasil
pedro.oliveira@geoexplore.com.br

Abstract. *This paper describes some essential aspects of a local Spatial Data Infrastructure (SDI) implementation. Initially, a brief literature review about Service Oriented Architecture (SOA) is developed. After, an analysis of SOA application possibilities is done, particularly in the case study: Belo Horizonte City Information and Informatics Company (Prodabel). Finally, some implementation alternatives of the proposed model are considered based on evolution scenarios.*

Resumo. *Este artigo descreve alguns aspectos cruciais da construção de uma Infra-Estrutura de Dados Espaciais (SDI) municipal. Faz-se, inicialmente, uma breve revisão teórica sobre Arquitetura Orientada por Serviços (SOA), seguida de uma análise sobre as possibilidades de sua aplicação específica à situação em estudo: Empresa de Informática e Informação do Município de Belo Horizonte - Prodabel. Por fim, avalia-se algumas alternativas de implementação do modelo proposto, considerando cenários de sua evolução.*

1. Contextualização – As Mudanças de Arquitetura em SIG

A arquitetura de aplicações geográficas tem acompanhado a evolução das características dos sistemas de informação em geral, com clara tendência ao uso do modelo orientado por serviços, baseado em plataforma web. As vantagens oferecidas por esse modelo podem ser entendidas comparando arquiteturas mais novas com as tradicionais, pouco flexíveis, e limitadas em relação ao conjunto de tecnologias hoje existentes. De fato, Sistemas de Informação Geográficos (SIG) fechados[1], baseados em uma camada ou em duas camadas, no modelo cliente/servidor e suas variações, predominaram no mercado por muito tempo [2]. Nos últimos anos, no entanto, com o desenvolvimento de arquiteturas aderentes aos padrões da rede mundial de computadores, tais características se tornaram inadequadas ao contexto Web. Assim, o mercado tem adotado, progressivamente, modelos de três camadas ou mais. A orientação por objetos foi outro ingrediente importante para essa mudança, permitindo modelar dados e aplicações com maior flexibilidade, voltadas ao reuso. Porém, a consolidação da modelagem de serviços como padrão arquitetural de mercado depende de uma série de questões - de negócio, culturais e tecnológicas, a serem atendidas. Nesse aspecto, a Arquitetura Orientada por Serviços (*Service Oriented Architecture* - SOA) se torna atraente como

[1] Fechados no sentido de serem totalmente construídos sobre tecnologias proprietárias e monolíticas.

uma tentativa de agregar valor ao negócio, permitindo modelar e aprimorar processos com redução do desperdício dos recursos de Tecnologia da Informação (TI). Em SIG, particularmente, SOA é atraente por permitir a concepção de sistemas interoperáveis e distribuídos, que funcionem utilizando módulos fracamente acoplados.

Essa visão coincide com algumas das propostas mais bem sucedidas de organizações normalizadoras e padronizadoras da área de geotecnologias, especialmente o *Open Geospatial Consortium* (OGC) [8]. Padrões de serviços geográficos, tais como *Web Feature Service* (WFS), *Web Map Service* (WMS) e, mais recentemente, *Web Processing Service* (WPS), dentre muitos outros, são exemplos de iniciativas nessa direção. A extensão desses padrões para formação de Infra-Estruturas de Dados Espaciais (*Spatial Data Infrastructure* – SDI) parece ser a solução mais viável para evoluir a arquitetura SIG na escala de países e regiões, como demonstram os esforços em desenvolvimento na comunidade européia (INSPIRE) [11] e outros.

Este trabalho propõe a aplicação dos conceitos de SOA e SDI por meio de um estudo de caso, em que um SIG de âmbito municipal seria totalmente remodelado, evoluindo para um SDI local (seção 4). Foi escolhido como serviço piloto, nesse contexto, um processo de negócio para geocodificação com endereçamento flexível [7], aplicável à prefeitura do município de Belo Horizonte (MG).

A próxima seção (2) apresenta aspectos conceituais de SOA, enquanto a seção 3 discute a arquitetura de uma SDI local. As seções 4 e 5, respectivamente, desenvolvem um estudo de caso e algumas conclusões (ainda parciais) deste trabalho.

2. Arquitetura Orientada por Serviços

Booch, Rumbaugh e Jacobson [1] conceituam a arquitetura de uma aplicação de maneira bem ampla. Ela seria composta por um conjunto de decisões importantes em questões como: organização do sistema, seleção de seus elementos estruturais e interfaces, considerando comportamento, decomposição desses elementos, e finalmente o estilo para composição dos elementos estáticos e dinâmicos, dentro do sistema. Fowler [5] considera que a arquitetura de uma aplicação possui dois objetivos essenciais: decompor esse sistema em suas partes principais, em alto nível, e representar um modelo geral de forma estável, ou seja, sem grande tendência a alterações.

A arquitetura orientada por serviços [9] pode ser entendida como uma proposta de separação lógica das camadas de uma aplicação, visando a componentização da lógica do negócio. Nesse contexto, a explicitação de uma camada de serviços é fundamental para permitir a interoperabilidade e o reuso. É importante destacar que, mesmo sendo uma condição essencial, serviços por si só não implementam SOA.

Embora existam diversas razões para se investir em SOA, os desafios na utilização de uma arquitetura tão inovadora e com poucos exemplos de uso corporativo (menos ainda no Brasil) são grandes. Academicamente observa-se um grande número de novos trabalhos sobre esse tema, cujas justificativas [7] seriam suas vantagens:

- Possibilitar o gerenciamento automatizado de processos de negócio;
- Facilitar a integração entre sistemas;
- Permitir o reuso de sistemas legados;
- Favorecer a adaptação de aplicações a mudanças tecnológicas.

Não se verifica a necessidade de padronização do modelo de dados das diferentes aplicações envolvidas, mas certamente a diversidade de modelos, formatos e tecnologias pode converter-se num desafio para uma solução particular. Tal situação se agrava ao manipular dados geográficos, considerando as dificuldades de conversão e de troca de dados de diferentes formatos, com sua semântica.

A seção 3 discute um modelo de SDI local, destacando o papel dos serviços.

3. Uma Arquitetura para Infra-Estrutura de Dados Espaciais Municipal

Davis e Alves [3] propuseram uma arquitetura para *Spatial Data Infrastructure* (SDI) de âmbito local, apresentada esquematicamente na figura 1.

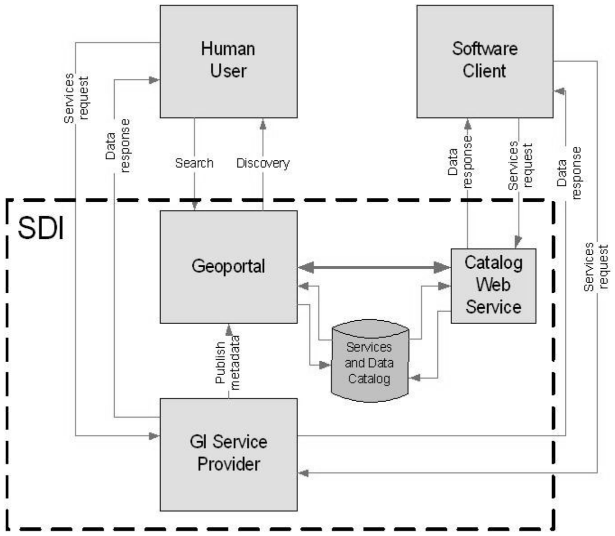


Figura 1: Geoportais e SDI para uso local (Fonte: Davis Júnior e Alves, 2006) [3]

A proposta de Davis e Alves baseia-se nos padrões OGC [4]. Os autores sugerem sua extensão para atender também aos requisitos do World Wide Web Consortium (W3C) [10]. A proposta de SDI se diferencia completamente do modelo de um SIG, transferindo o foco do sistema, anteriormente nos dados, para os serviços. Isto pode ser observado na próxima seção (4), através de um estudo de caso.

4. Estudo de Caso – O SDI do Município de Belo Horizonte

Belo Horizonte é uma capital de estado que adota, desde o início da década de 1990, soluções inovadoras e desafiadoras para a estruturação de seu Sistema de Informação Geográfico. O desenvolvimento de aplicativos e a customização do SIG, a cargo da Empresa de Informática e Informação do Município de Belo Horizonte (Prodabel), sempre foi visto pela equipe de geoprocessamento como um desafio profissional. Nesse contexto, formou-se na empresa um grupo de pesquisadores que foi buscar na academia conhecimentos técnicos e científicos que pudessem reverter em soluções práticas.

Ao longo de anos de trabalho e investimentos, a Prodabel tornou-se a detentora do Mapa Urbano Básico (MUB) local, além de um conjunto de dados temáticos variados, de características *sui generis*. Isto ocorreu porque, além de ter de lidar com um grande número de camadas de informação relacionadas à sua área de atuação [6] a empresa lidera um convênio de cooperação técnica (Grupo de Gestão da Informação –

GGI) voltado para o intercâmbio de dados entre diversos órgãos públicos municipais, estaduais e federais. Cumpre destacar que, neste caso, a opção por um modelo baseado em SDI se deu pela necessidade de evoluir no uso de novas tecnologias e ampliar a base de dados, acrescentando gradativamente novas categorias e camadas oriundas de outros órgãos da prefeitura e do convênio (GGI).

A situação atual do SIG existente na Prodabel é apresentada na figura 2:

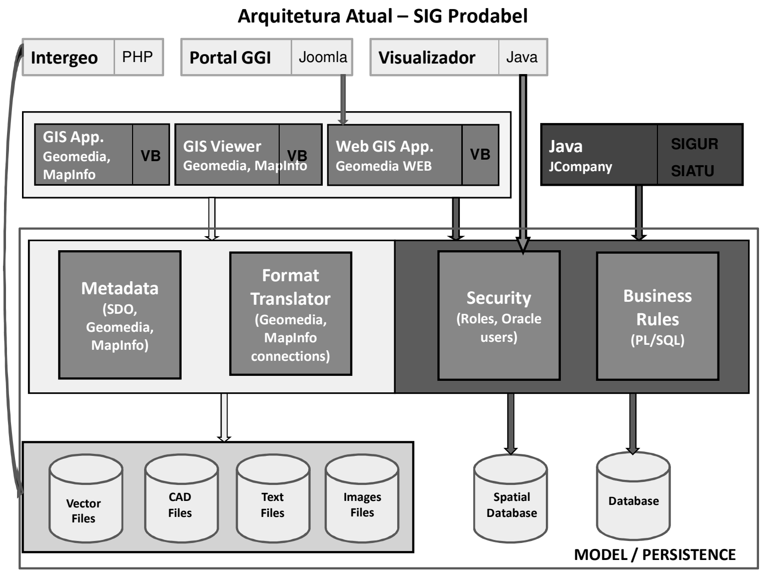


Figura 2: Arquitetura Atual do SIG. Fonte: Prodabel, 2008

Na figura 2 pode-se observar a miscelânea de tecnologias atualmente em uso na manutenção dos mais de cem componentes do Cadastro Técnico Municipal (CTM), e na obtenção de informações. As ferramentas utilizadas na manutenção da base de dados são produtos comerciais. Para visualização e consultas foram desenvolvidas internamente aplicações (*applets* e *servlet*s, padrão JEE), além do uso de ferramentas livres. Como há muitas formas de acesso às informações, optou-se por definir os mecanismos de segurança de acesso e as regras de negócio na camada de modelo / persistência (*scripts* SQL). Outros recursos disponíveis nessa camada são estruturas de metadados (criadas e mantidas pelos *softwares* e pelo SGBD) e ferramentas de tradução (nativas dos produtos), responsáveis pela interoperabilidade no acesso aos múltiplos formatos suportados (SIG, *Computer Aided Design* - CAD, imagem, texto).

A mudança de configuração do SIG da Prodabel, iniciada em 2001 com o projeto "Prospecção de Ferramentas SIG", culminou numa definição de arquitetura apoiada em três pilares: base de dados integrada, interoperabilidade e acesso descentralizado. Essas linhas mestras ditaram a escolha de ferramentas, tornando possível a integração do sistema à Rede Municipal de Informática (RMI). Assim, o modelo da figura 2 pode ser considerado uma etapa preparatória para o SDI (figura 3).

A figura 3 apresenta o novo modelo de arquitetura da Prodabel, que visa atender não somente à realidade da evolução observada em geoprocessamento na empresa como também propiciar o intercâmbio de dados com outros órgãos e a publicação dessas informações em ambiente Web, tudo isso integrado numa arquitetura de serviços.

Arquitetura proposta – SDI Prodabel

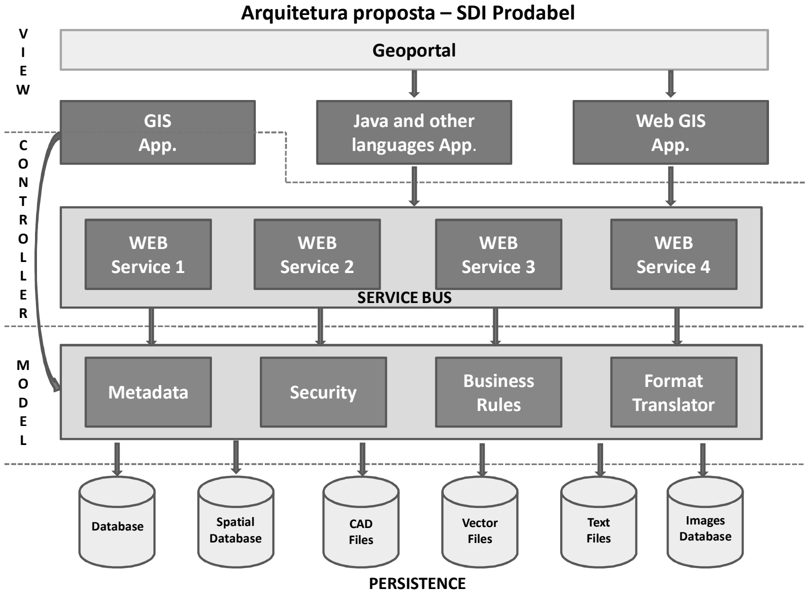


Figura 3: Arquitetura Proposta para SDI - Belo Horizonte. Fonte: Prodabel, 2008

A figura 3 mostra que, no novo modelo, aplicações SIG comerciais, já em uso na organização, não serão descartadas. Porém, dada sua característica cliente/servidor, elas farão acesso diretamente à camada de modelo, que incorpora regras de negócio para acesso às diferentes bases de dados. A adoção do padrão *Model, View, Controller* (MVC) atende às características de segmentação vertical, enquanto a construção de um *middleware* de serviços (segmentação horizontal) atende às práticas de SOA. Por sua vez, aplicações como *Applets* e *Servlets* Java e ferramentas para visualização de mapas e análise espacial na Web podem interagir diretamente com os serviços da camada de controle. Um exemplo de serviço a ser disponibilizado é a geocodificação de endereços [4], [7], que permite encontrar um endereço com alto grau de precisão. Esse serviço é composto de três etapas: *Parsing*, *Matching*, *Locating*. Sua modelagem e implementação são apresentadas em [7]. A escolha desse primeiro serviço deveu-se à realização de pesquisas e testes bem sucedidos em ambiente controlado. Outros serviços serão progressivamente agregados ao *service bus*, como mostra a figura 3.

A camada de modelo, na figura 3, certamente merece uma atenção especial. Visando a separação entre os formatos de armazenamento e os de troca, ao levar em consideração a existência de esquemas de metadados necessários para o acesso às informações no formato que se deseja, optou-se por separar o modelo semântico das estruturas de dados (*persistence*). Assim, regras de segurança e de integridade gerais serão desenvolvidas preferencialmente em linguagem não proprietária do banco, prevalecendo sobre todas as formas de acesso.

5. Considerações Finais

Pretendeu-se, nesta breve análise, descrever algumas questões cruciais para atender às necessidades dos usuários de um SDI local. A primeira delas é definir um Geoportal como meio de entrada para o sistema. A segunda é a possibilidade de sistemas Web poderem ler e tratar informações alfanuméricas associadas a camadas geográficas, sem comprometer a integridade dessas. A terceira questão diz respeito à interoperabilidade entre formatos de dados geográficos, que pode ser tratada por padrões de metadados aliados a um serviço tradutor de formato capaz de quebrar a complexidade inerente a essas conversões. O quarto desafio é conseguir que esse modelo tenha escalabilidade e possa, assim, evoluir junto com a demanda. Outra questão é manter o modelo de dados (alfanuméricos, vetoriais, imagens) íntegro e disponível para as diferentes classes de usuários: geradores e consumidores de informação. Esses requisitos estão condicionados a questões tecnológicas não triviais, condicionadas à pesquisa. Nesse caminho, a adoção de bons padrões de projeto e de tecnologias abertas, aderentes aos padrões OGC para serviços e armazenamento, pode ser fundamental.

Espera-se, a partir deste ponto, desenvolver uma prova de conceito que valide a hipótese de interoperabilidade do modelo. Posteriormente, poderá ser conduzida uma revisão do modelo proposto, agregando as boas práticas avaliadas e descartando as ineficazes.

Finalmente, por meio de estudos e pesquisas, pretende-se buscar ainda a validação de partes do modelo capazes de gerar inovação em áreas desafiadoras e promissoras como:

- Serviços padrão OGC adequados às necessidades dos diferentes tipos de usuários;
- Comparação entre padrões de serviços OGC e W3C para escolha de soluções específicas;
- Testes de desempenho com a nova arquitetura (em ambiente simulado e real);
- Desenvolvimento e uso de um serviço para compartilhamento de dados geográficos;
- Avaliação da SDI para aplicação ao amplo conjunto de informações dos diversos órgãos envolvidos, com seus diferentes modelos de dados e temas.

Referências

[1] Booch, G; Rumbaugh, J; Jacobson, I. (2000) UML, Guia do Usuário. Elsevier, Rio de Janeiro.

[2] Davis Jr., C. A.; Oliveira, P. A. (2002) SIG Interoperável e Distribuído para Administrações Municipais de Grande Porte. Informática Pública, Belo Horizonte (MG), v. 4, n. 1, p. 121-141.

[3] Davis Jr., C. A. D.; Alves, L. L. (2006) *Interoperability through Web Services: Evaluating OGC Standards in Client Development for Spatial Data Infrastructures*. In: *Proceedings of Brazilian Symposium on GeoInformatics* (Geoinfo, 2006).

[4] Davis Jr., C. A.; Fonseca, F. T. (2007) *Assessing the Certainty of Locations Produced by an Address Geocoding System*. Geoinformatica 11(1), 103-129, 2007.

[5] Fowler, M. (2006) Padrões de Arquitetura de Aplicações Corporativas. Bookman, São Paulo.

[6] Oliveira, P. A.; Oliveira, M. P. (2005) Usos de um Sistema de Informação Geográfico (SIG) em Cadastro Técnico Municipal (CTM): A experiência de Belo Horizonte. In: Informática Pública, ano 8, n.1.

[7] Oliveira, P. F. A.; Oliveira, P. A. (2008) Arquitetura Orientada a Serviços (SOA): Um Estudo de Caso em Sistemas de Informação Geográficos. In: Anais SIMGEO 2008.

[8] www.ogc.org

[9] www.soainstitute.org/index.php

[10] http://www.w3.org/

[11] http://inspire.jrc.ec.europa.eu/

Sharing executable models through an Open Architecture based on Geospatial Web Services: a Case Study in Biodiversity Modelling

Karla Donato Fook[1,2], Silvana Amaral[1], Antônio Miguel Vieira Monteiro[1], Gilberto Câmara[1], Marco A. Casanova[3]

[1]Image Processing Division – National Institute of Space Research (INPE)
São José dos Campos – SP – Brazil
[2]Computer Science Department
[2]Centro Federal de Educação Tecnológica do Maranhão (CEFET–MA)
São Luís – MA – Brazil
[3]Informatics Department – PUC-Rio
Rio de Janeiro – RJ – Brazil

{karla, silvana, miguel, gilberto}@dpi.inpe.br,

casanova@inf.puc-rio.br

***Abstract**. Biodiversity researchers develop predictive models for species occurrence and distribution which are useful for biodiversity conservation policies. Species distribution modelling tools need to locate and access large amount of data in different sources and produces results from different algorithms. In this scenario, collaboration is an essential feature to improve this research area. Scientists need to share models, data and results to get new discoveries. This paper presents advances in Web Biodiversity Collaborative Modelling Services (WBCMS) development. These services support sharing of modelling results and information about its generation. WBCMS also enable researcher to make new experiments based in previous one. Scientists can use WBCMS to compare experiments, to make new inferences and to improve their studies. A case study explains the model instance usage.*

1. Introduction

Biodiversity research uses tools that allow performing inferences about diversity and abundance of species in different areas. Species distribution models combine *in situ* species data with geographical data. Their results support biodiversity protection policies, are useful to forecast of the impacts of climate change, and help detect problems related to invasive species. Since such data sets may be archived by different institutions, the scientist needs to locate the data sets and make them interoperate. These points create challenges that lead to data representation, management, storage, and access problems. In addition, the scientist would like to share his experiments results with the community and compare it with similar work done elsewhere.

This scenario points to the need for a computational infrastructure that supports collaborative biodiversity studies, allowing sharing data, models and results [Ramamurthy 2006]. Each of these three aspects needs a different strategy. Sharing data needs information about the location of repositories and archival formats. Sharing models needs understanding about the applicability of each algorithm to the species being modelled; it also requires a good documentation explicit and implicit assumptions behind the model. Sharing results needs communication of the species distribution maps as well as producing reports and adding comments. In this context, metadata are useful in order to disambiguate the data and enable reuse. One kind of metadata is provenance, which records data about scientific experiments [Simmhan, Plale and Gannon 2005]. Provenance metadata allows researchers to capture relevant information about scientific experiments, and to assess the experiment quality and timeliness of results [Greenwood, Goble, Stevens et al. 2003; Marins, Casanova, Furtado et al. 2007].

This paper reports advances on development of the Web Biodiversity Collaborative Modelling Services (WBCMS). They are geospatial web services that support cooperation on a species distribution modelling network, including sharing modelling results and its provenance, and enabling researchers to perform new experiments based in previous ones. Prototypes were implemented. An early prototype stored algorithms information in the database and does not produce the model instance. For more details, see [Fook, Monteiro and Câmara 2007]. A new prototype was developed. This prototype is more robust than the early prototype. The main differences between WBCMS prototypes are that the current prototype composes the model instance, and also enable researcher to reuse model instance data to produce new experiments. The WBCMS architecture is part of the OpenModeller[1] Project, a framework for collaborative building of biodiversity models [Muñoz 2004; Giovanni 2005; OpenModeller 2005].

This work is organized as follows. Section 2 presents related work. Section 3 describes WBCMS in detail. In Section 4, we show the current prototype by model instance use example. Finally, section 5 presents final remarks and further work.

[1] http://openmodeller.cria.org.br/

2. Related Work

Trends have enabled a new generation of data services in the scientific community. Web services stand out to support distributed applications in geospatial domain, where geographical application are divided in a series of tasks, organized in a workflow. Bernard et al. (2003) have developed a “road blockage” service, which solve more complex tasks by static chaining several simple services. WS-GIS approach is an SOA-based SDI which aims to integrate, locate, and catalog distributed spatial data sources [Leite-Jr, Baptista, Silva et al. 2007].

The *GeoCatalog* is a tool that implements a software architecture for automated geographic metadata annotation generation [Leme, Brauner, Casanova et al. 2007]. Díaz et al. (2007) designed a gvSIG[2] extension to collect automatically metadata. This application aids users to publish imagery or cartographic data in a Spatial Data Infrastructure. The Earth System Science Workbench (ESSW) is a metadata management and data storage system for earth science researchers. Their infrastructure captures and keeps provenance information for proving credibility of investigator-generated data [Frew and Bose 2001].

In biodiversity field, Best et al. (2007) use geospatial web services to automate the scientific workflow process in marine mammal observations from OBIS-SEAMAP[3]. Web Service Multimodal Tools for Biodiversity Research, Assessment and Monitoring Project (WeBIOS) provides scientists with a system that supports exploratory multimodal queries over heterogeneous biodiversity data sources [WeBios 2005]. BioWired project proposes a P2P grid architecture that supports biodiversity data access by large distributed database [Alvarez, Smukler and Vaisman 2005]. BiodiversityWorld project intends to make available heterogeneous data sources and biodiversity analytic tools in a Grid [Jones, White, Pittas et al. 2003; Pahwa, White, Jones et al. 2006].

The approaches above aim to integrate and share geographical data and tools. However, they do not aim to share modelling results. Our approach aims to support sharing descriptive information about spatial data, and relevant information objects. Our goals are to publish modelling experiments and their provenance, to make it available

[2] www.gvsig.gva.es

into catalogues, and to enable researchers to perform new models based in catalogued model instances.

3. WBCMS description

This section presents the Web Biodiversity Collaborative Modelling Services (WBCMS), a set of geospatial Web services that enables sharing of modelling experiments, and reusing of these data in new experiments.

This approach aims to capture the explicit and implicit information inserted in a biodiversity experiment, in our case, a species distribution modelling. A key idea behind WBCMS is a model instance. It includes data and metadata related to models, results and algorithms and describes an experiment as a whole. The idea is that the researcher examines model instances and be able to understand how a result was produced. He can then compare experiment results to reproduce them, and to use them for his own models. He can get answers for queries such as *"What species are being modelled?"*, *"Where does the data come from?"*, *"Which environmental variables are used?"*, and *"If I have a question, how can I look for similar results?"*. So, consider a distributed environment in which researchers perform species distribution modelling locally, and wish to share their experiments (Figure 1).

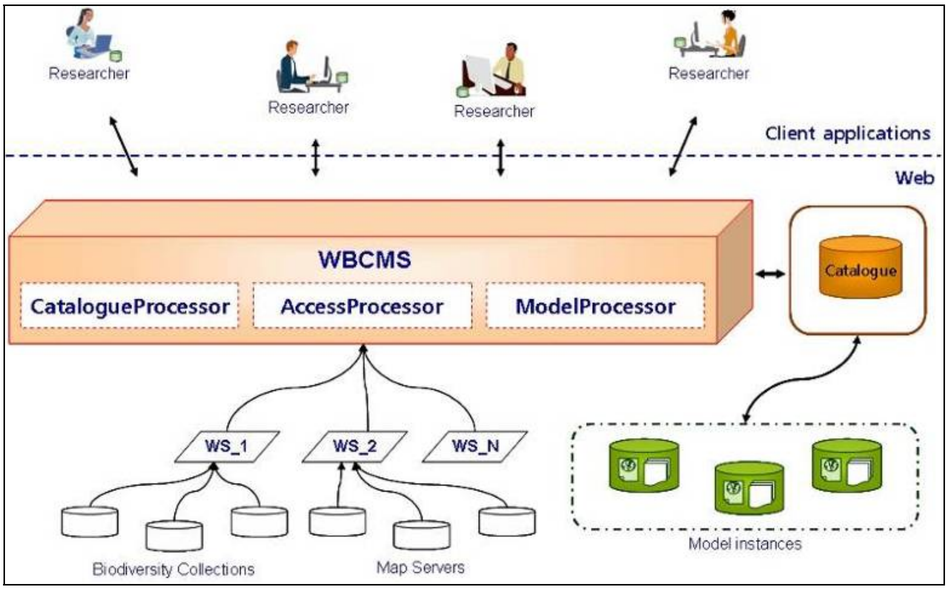


Figure 1. WBCMS Architecture

[3] http://seamap.env.duke.edu

Briefly, researchers can use the WBCMS to (a) share their modelling experiment, (b) access and evaluate experiments, and (c) perform new models based in catalogued models. Therefore, WBCMS builds a catalogue of model instances and holds processors to handles with each activity: *Catalogue Processor*, *Query Processor*, and *Model Processor*. These processors include a set of web and geoweb services. The model instance catalogues can be in different institutions and holds information related to different kind of model, such as environmental and urban models. Therefore, one challenge in this approach is to specify the model instance, since it must provide researchers with the necessary information for a better understanding of an experiment. We present our idea of a model instance in next subsection.

3.1 Model instance outline

This subsection describes the model instance in WBCMS architecture. It aims to describe a modelling experiment as whole. The model instance idea includes several types of models such as Land Use and Coverage Change, and Natural Hazards models. In our case, we are working with species distribution modelling where the modeled object is a species. The model instance includes data and metadata about the model, its generation process, and experiment results (see Figure 2).

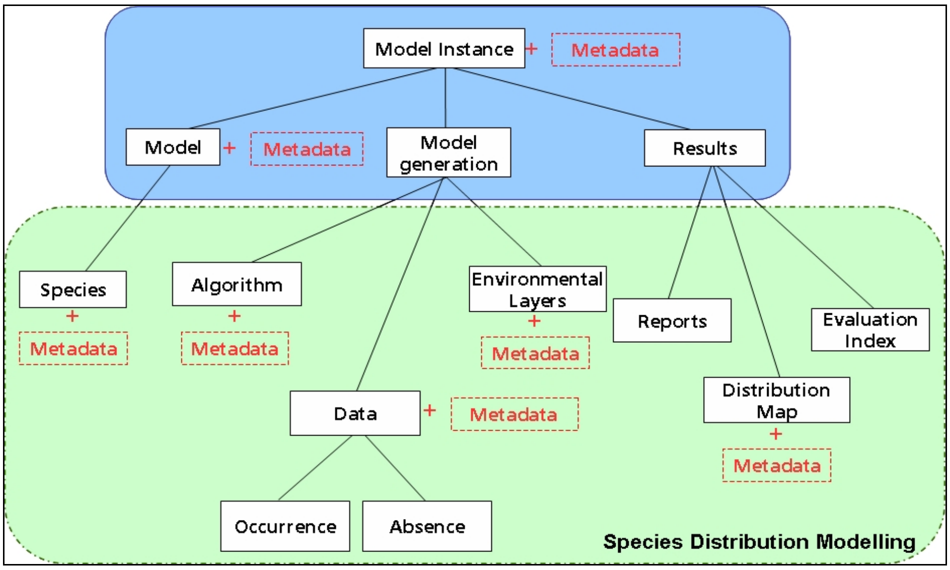


Figure 2. Model Instance Diagram

Figure 2 shows the model instances diagram and highlights that each element of model instance contains their own metadata. The model instance includes information as shown in Table 1.

Table 1. **Model instance elements**

Element	Desccription	Information
Model instance	Global information related to modelling experiment, and researcher's notes that help scientists in experiment analysis.	Model instance name, title, description, author, affiliation, creation date, running time, and modelling motivation (or question) comments, confidence degree and its justification
Model	Information about the used model, and information related to modelled object, in this case, modelled species.	Model author, description and version. Species: taxonomic data (classification), and their metadata (author, status, online resource, and reference date).
Model generation	Input data and used algorithm, as well as metadata such as execution time and messages.	Species occurrence and absence points (latitude and longitude), and environmental layers are input data examples. It also includes algorithm parameters and metadata like description, version, author, and contact.
Results	Set of modelling result files.	Reports, georeferenced maps, and model evaluation indexes.

Besides metadata about experiment results, a model instance includes other information such as species taxonomic data (see Table 1). Species-occurrence presents different reliability degree to biodiversity researchers, because these records have different sources and methods. Therefore, make it available is not enough to assure their use by the community. The minimum requirements for a species occurrence record are its geographical positioning, and its taxonomic identification together with metadata such when, and details about where the specimen was collected [Guralnick, Hill and Lane 2007].

We used the ISO19115 standard [ISO 2003] to describe the model instance. It includes the experiment provenance, and provides evaluation features for accessing the experiment. The model instance has a set of metadata to describe itself globally, and to describe model instance elements. Therefore, there are metadata copies to different components, for instance use the reference date metadata to points to different dates:

experiment performing, experiment cataloguing and species data recovering. WBCMS extracts part of metadata from result files and recovers another part from web. On the other hand, the researcher needs to inform extra metadata related to experiment in client application, as description and lineage. Next subsection describes WBCMS processors in detail.

3.2 WBCMS Processors

This subsection presents the WBCMS processors. The WBCMS *Catalogue Processor* publishes a model instance. The researcher uses a catalogue application to send basic experiment elements to WBCMS. The *Catalogue Processor* receives modelling result data, accesses remote data, and composes model instance. Then, the WBCMS inserts a model instance into the repository. Figure 3 details the WBCMS *Catalogue Processor*.

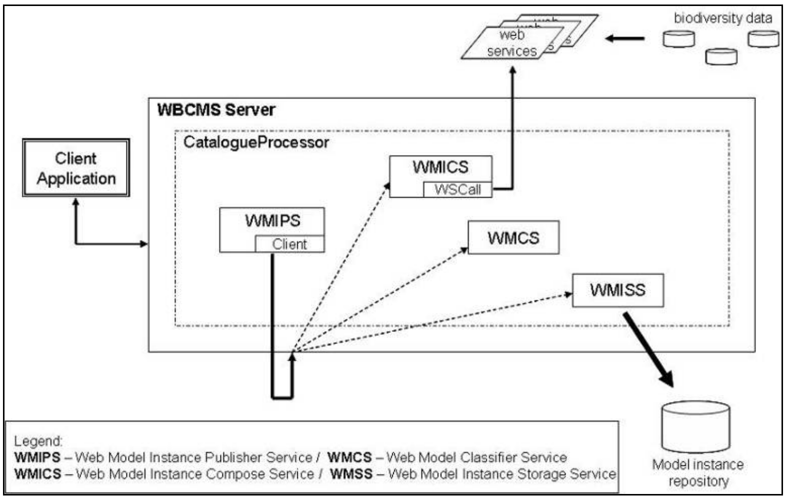


Figure 3. WBCMS Catalogue Processor

The *Catalogue Processor* includes following services: WMIPS – Web Model Instance Publisher Service, WMICS – Web Model Instance Compose Service, WMCS – Web Model Classifier Service, and WMISS – Web Model Instance Storage Service. The WMIPS is an orchestration service that controls the other catalogue processor services. WMCS uses model metadata to perform a model instance classification. WMICS recovers biodiversity data and metadata from web to complement the model instance. Finally, WMISS inserts a model instance into a repository.

A researcher uses the WBCMS *Access Processor* to retrieve model instances. This processor uses the OGC WFS – Web Feature Service [OGC 2005] and two services: WMIQS – Web Model Instance Query Service, and WMIRS – Web Model Instance Retrieval Service (see Figure 4).

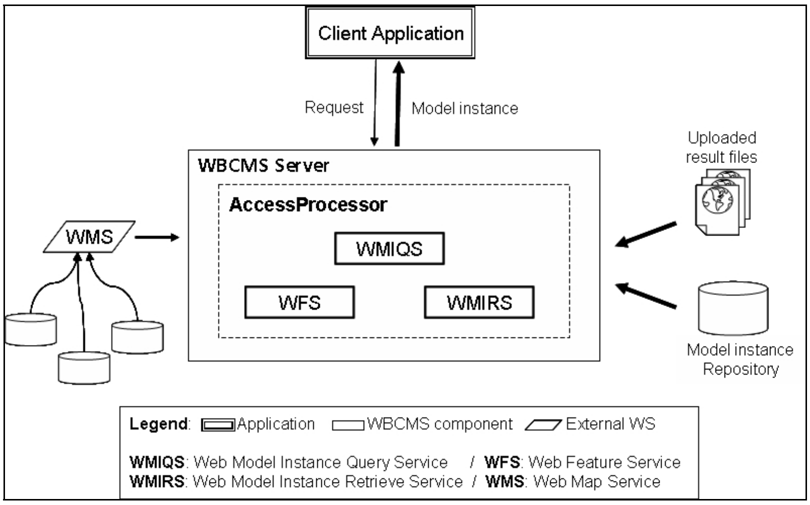


Figure 4. WBCMS *Access Processor*

The WBCMS *Access Processor* receives requests from a client application and uses WMIQS to handle queries, and WMIRS and WMS [OGC 2006] to recover the model instance, and make it available.

The researchers can reuse catalogued data to execute remotely new models using the WBCMS *Model Processor*. This processor includes the WMRS – Web Model Run Service, and uses the OMWS – OpenModeller Web Service. The WMRS is responsible to: (a) prepare input data and allows user to change algorithm parameters, (b) call OMWS to perform the new model, and (c) increment the model instance run count at each model instance reuse. The last activity allows a statistic evaluation of the instance model reuse. We use the UML communication diagram to show the WBCMS *Model Processor* usage (Figure 5).

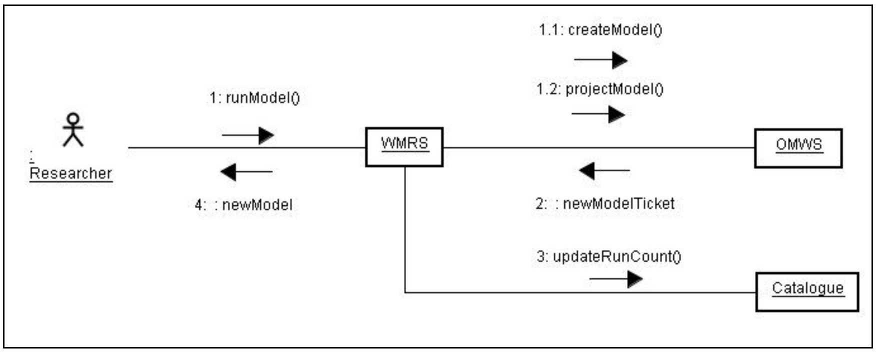


Figure 5. WBCMS *Model Processor*

The WMRS receives the researcher request to perform a new model, and send the necessary request and input data to OWMS produce it. The OMWS receives occurrence data and algorithm parameters from client, performs the model, and returns the produced species distribution model [Giovanni 2005; Sutton, Giovanni and Siqueira 2007]. We developed a prototype as proof of concept of our approach. Figure below shows WBCMS class diagram.

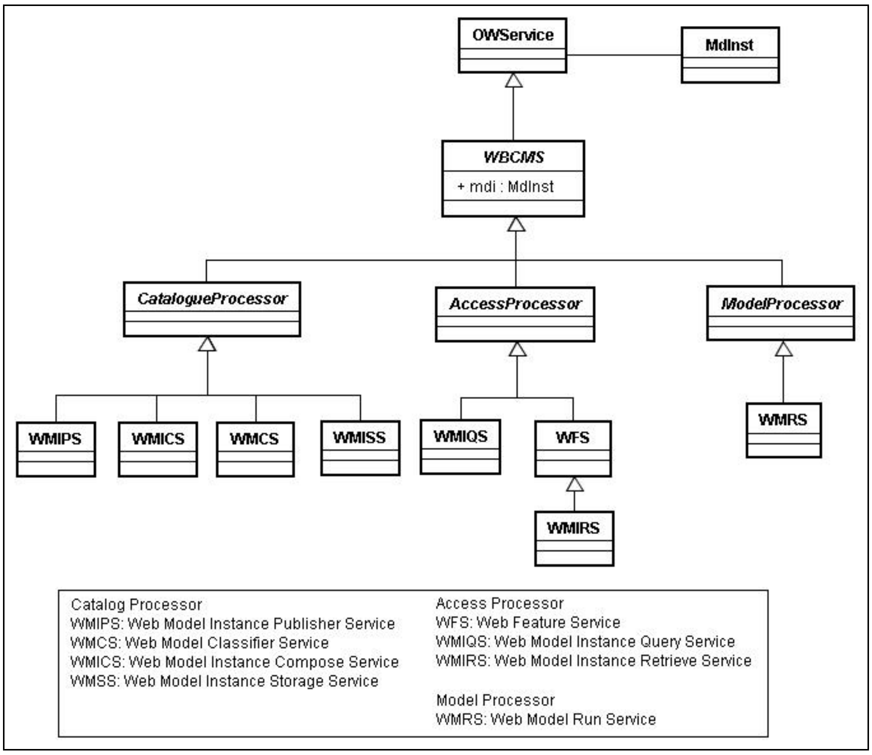


Figure 6. WBCMS Class Diagram

Figure 6 shows web and geoweb services of proposed architecture. There is an association relation between WBCMS and MdInst (*Model Instance*) classes. Next section presents an example of the WBCMS prototype functionalities.

4. WBCMS Prototype: A model instance usage example

This section presents an example that shows how the WBCMS makes a model instance available and how a researcher can produce new species distribution models. The example considers the *Coccocypselum erythrocephalum* Cham. & Schltdl. Species. The genus *Coccocypselum* belongs to Rubiaceae family, one of the most important families in the tropics.

In this example, we show the model instance **md_CErythr**. Initially, the researcher uses the OpenModeller Desktop [Giovanni 2005; Sutton, Giovanni and Siqueira 2007] to produces the species distribution model. This model consists of several result files, such as distribution map, reports and configuration files. The researcher uses the *Model Instance Catalogue* application to capture provenance information from result files, to inform personal comments about the experiment, and to publish the model instance into the catalogue.

The researcher can access **md_CErythr** model instance using the *Model Instance Access* application. This application enables the scientist to visualize each model instance element, and to perform new models based in previous ones. Figure 7 illustrates the modelling results visualization.

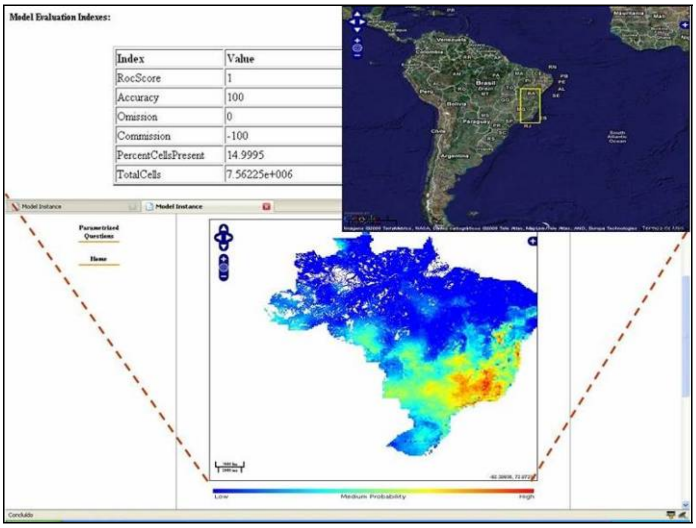


Figure 7. Model instance access application – Modelling result

Besides **md_CErythr**'s species distribution map, the form displays model evaluation indexes, and map with bounding box showing the species area (Figure 7). The evaluation indexes and author comments about the experiment help the user to capture relevant aspects of the species distribution model. The *Model Instance Access* application also makes available metadata about algorithms and model instance authors. The researcher can use WBCMS to perform new models reusing catalogued model instance data. Figure 8 displays **md_CErythr**'s algorithm metadata and parameters. The researcher can change algorithm parameters and layers to produce different models using OMWS (OpenModeller Web Service). So, several models can be produced (Figure 9).

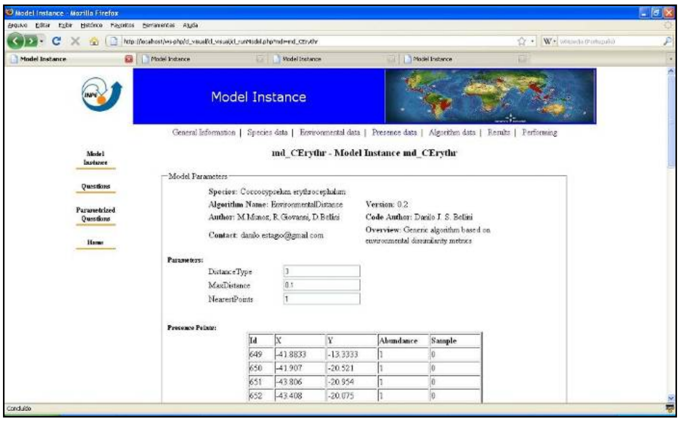


Figure 8. md_CErythr model instance reuse form

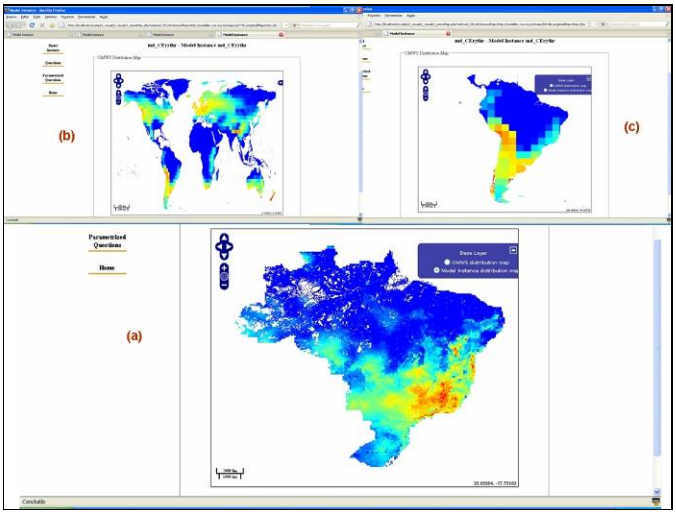


Figure 9. Species distribution maps based on md_CErythr model instance

Figure 9 shows the **md_CErythr** species distribution map (a), and distributions maps based in this model instance (b, and c labels). Algorithm parameters and layers were changed to produce these models. A detailed discussion of distribution maps analysis is beyond the scope of this paper. The objective is enables scientist to compare different distribution maps and to make new inferences about his studies.

5. Final Remarks

We presented in this paper advances in the development of Web Biodiversity Collaborative Modelling Services (WBCMS), a set of geospatial web services that aim at making it available modelling experiment results in a species distribution network, and enable researchers to perform new models based in previous ones.

We introduced the model instance idea that aims at describing an experiment as whole. Then, a set of ISO metadata elements were selected to describe a model instance.

We used compliant OGC web services to show model instance elements. However, existent specifications are not enough to work with the sharing of model description and results. In addition, we developed web services to handle with model instance complexity. We also included in the paper a model instance example illustrating the WBCMS use.

Our experiments, have demonstrated the validity of the proposals and ideas presented in this paper. We consider this line of work promising, even though more tests with a larger volume of modelling experiments are required. Finally, we remark that we will to improve WBCMS to handle more complex query predicates, and to provide model instance reuse statistics.

Acknowledge

Special thanks go to Dr. Cristina Bestetti Costa for their relevant comments and species occurrence data. We also thanked OpenModeller Project (FAPESP process: 04/11012-0), and FAPEMA[4] (In Portuguese: Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão) for partially supporting this research.

References

Alvarez, D., Smukler, A. and Vaisman, A. A. (2005). "Peer-To-Peer Databases for e-Science: a Biodiversity Case Study." Proceedings 20th Brazilian Symposium on Databases and 19th Brazilian Symposium on Software Engineering.

Bernard, L., Einspanier, U., Lutz, M., et al. (2003). Interoperability in GI Service Chains-The Way Forward. 6th AGILE Conference on Geographic Information Science, Muenster.

Best, B. D., Halpin, P. N., Fujioka, E., et al. (2007). "Geospatial web services within a scientific workflow: Predicting marine mammal habitats in a dynamic environment." Ecological Informatics **2**: 210-223.

Díaz, L., Martín, C., Gould, M., et al. (2007). Semi-automatic Metadata Extraction from Imagery and Cartographic data. Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS 2007), Barcelona (Spain).

Fook, K. D., Monteiro, A. M. V. and Câmara, G. (2007). Web Service for Cooperation in Biodiversity Modeling. Advances in Geoinformatics. C. Davis and A. M. V. Monteiro, Springer**:** 203-216.

Frew, J. and Bose, R. (2001). Earth System Science Workbench: A Data Management Infrastructure for Earth Science Products. 13th International Conference on Scientific and Statistical Database Management (SSDBM), Virginia, USA.

Giovanni, R. D. (2005). The OpenModeller project. BiodiversityWorld GRID workshop, Edinburgh, e-Science Institute.

[4] http://www.fapema.br/

Greenwood, M., Goble, C., Stevens, R., et al. (2003). Provenance of e-Science Experiments - experience from Bioinformatics. 2nd UK e-Science All Hands Meeting, Nottingham, UK.

Guralnick, R. P., Hill, A. W. and Lane, M. (2007). "Towards a collaborative, global infrastructure for biodiversity assessment." Ecology Letters **10**: 663-672.

ISO (2003). ISO 19115 Geographic Information - Metadata. Geneva, International Organization for Standardization (ISO).

Jones, A. C., White, R. J., Pittas, N., et al. (2003). BiodiversityWorld: An architecture for an extensible virtual laboratory for analysing biodiversity patterns. UK e-Science All Hands Meeting, Nottingham, UK, Cox, S.J.

Leite-Jr, F. L., Baptista, C. S., Silva, P. A., et al. (2007). WS-GIS: Towards a SOA-Based SDI Federation. Advances in Geoinformatics. C. Davis and A. M. V. Monteiro, Springer**:** 247-264.

Leme, L. A. P., Brauner, D. F., Casanova, M. A., et al. (2007). A Software Architecture for Automated Geographic Metadata Annotation Generation. 2007 e-Science Workshop.

Marins, A., Casanova, M. A., Furtado, A., et al. (2007). Modeling Provenance for Semantic Desktop Applications. SEMISH - XXXIV Seminário Integrado de Software e Hardware, Rio de Janeiro, RJ, SBC.

Muñoz, M. (2004). openModeller: A framework for biological/environmental modelling. Inter-American Workshop on Environmental Data Access, Campinas, SP. Brazil.

OGC. (2005). "OpenGIS Web Feature Service (WFS) Implementation Specification." Panagiotis A. Vretanos. from http://portal.opengeospatial.org/modules/admin/license_agreement.php?suppressHeaders=0&access_license_id=3&target=http://portal.opengeospatial.org/files/index.php?artifact_id=8339

OGC (2006). OpenGIS Web Map Server Implementation Specification, OGC - Open Geospatial Consortium Inc.

OpenModeller. (2005). "openModeller: Static Spatial Distribution Modelling Tool." Retrieved agosto/2005, from http://openmodeller.cria.org.br/.

Pahwa, J. S., White, R. J., Jones, A. C., et al. (2006). Accessing Biodiversity Resources in Computational Environments from Workflow Applications. The Workshop on Workflows in Support of Large-Scale Science.

Ramamurthy, M. K. (2006). "A new generation of cyberinfrastructure and data services for earth system science education and research." Advances in Geosciences **8**: 69-78.

Simmhan, Y. L., Plale, B. and Gannon, D. (2005). "A survey of data provenance in e-Science." SIGMOD Record **34**(3): 31-36.

Sutton, T., Giovanni, R. d. and Siqueira, M. F. d. (2007). "Introducing openModeller - A fundamental niche modelling framework." OSGeo Journal **1**.

WeBios (2005). WeBios: Web Service Multimodal Tools for Strategic Biodiversity Research, Assessment and Monitoring Project, http://www.lis.ic.unicamp.br/projects/webios.

Discovering Location Indicators of Toponyms from News to Improve Gazetteer-Based Geo-Referencing

Cleber Gouvêa[1], Stanley Loh[1,2], Luís Fernando Fortes Garcia[2,3], Evandro Brasil da Fonseca[1], Igor Wendt[1]

[1]Centro Politécnico – Universidade Católica de Pelotas (UCPEL)
96010-000 – Pelotas – RS – Brasil

[2]Centro de Tecnologia e Computação – Universidade Luterana do Brasil (ULBRA)
92425-900 – Canoas – RS – Brazil

[3]Curso de Informática – Faculdade Dom Bosco
Porto Alegre – RS – Brasil

cleber AT sindiq.com.br, sloh AT terra.com.br, luis AT garcia.pro.br, {evandrofonseca08, igor.wendt} AT gmail.com

Abstract. *This paper presents an approach that identifies Location Indicators related to geographical locations, by analyzing texts of news published in the Web. The goal is to semi-automatically create Gazetteers with the identified relations and then perform geo-referencing of news. Location Indicators include non-geographical entities that are dynamic and may change along the time. The use of news published in the Web is a useful way to discover Location Indicators, covering a great number of locations and maintaining detailed information about each location. Different training news corpora are compared for the creation of Gazetteers and evaluated by their ability to correctly identify cities in texts of news.*

1 Introduction

Geo-referencing of texts, that is, the identification of the geographical context of texts is becoming popular in the web due to the high demand for geographical information [Sanderson and Kohler 2004] and due to the raising of services for query and retrieval like Google Earth (geobrowsers). The main challenge is to relate texts to geographical locations. However, some ambiguities may arise [Clough et al. 2004]:

- *Reference Ambiguity*: the same location may be referenced by many names;
- *Referent Ambiguity*: the same term may be used to reference different locations (for example, two cities with the same name);
- *Referent Class Ambiguity:* the same term may be used to reference different kinds of locations (for example, a street and a city with the same name).

For solving ambiguity problems, one of the alternatives is to utilize Word Sense Disambiguation techniques [Li et al. 2003] where co-occurrence of terms or collocations are identified in training texts. These associations are stored in structures called Gazetteers, that relate locations and references such as names of geographical entities, the kind of location (city, street, state, etc.), synonyms and also geographical coordinates [Hill et al. 2000].

Although the existence of predefined Gazetteers like Geonames [Geonames 2008] and Getty Thesaurus of Geographic Names [Tgn 2008] , they fail in coverage, lacking information about some countries, and they also fail by weak specialization, lacking detailed references to locations (fine granularity) as for example names of streets, squares, monuments, rivers, neighborhoods, etc [Leveling et al. 2006]. This last kind acts as indirect references to geographical locations and is important because texts about locations frequently utilize this kind of information instead of the explicit name of the location (for example, textual news in the Web). [Leveling and Hartrumpf 2007] call this kind of information as "Location Indicators" and identified some types:

- Adjectives: Rio de Janeiro ➔ "Wonderful city";
- Synonyms: Rio de Janeiro ➔ Rio;
- Codes and acronyms: BSB as the acronym for the airport of the city of Brasília;
- Idiom variations: São Paulo and San Pablo;
- Other geographical entities: names of authorities, highways, squares, airports, etc.

Location Indicators include non-geographical entities, like very important people related to the location, historical events or even temporary situations. The problem is that most of these indirect references are dynamic and may change along the time, and for this reason they do not appear in traditional Gazetteers, because pre-defined Gazetteers are manually created and maintained by people. According to [Delboni et al. 2007], the quality of Gazetteers depend on regular updates (global Gazetteers suffer about 20 thousand modifications per mouth [Leidner, 2004]). For this reason, these Gazetteers do not cover a great number of locations neither have much specific information.

The current work proposes the automatic creation of Gazetteers as a way to cover a great number of locations and for maintaining detailed information about each location. The idea is to utilize news published in the Web to generate and maintain a Gazetteer with detailed information, including indirect references (Location Indicators). Although the existence of works that utilize automatic techniques for supervised learning, these works usually demand manual annotation of the training corpus and are applicable only in specific idioms.

The approach proposed in this paper extracts Location Indicators from news, based on co-occurrence of proper names, without manual annotation and for whatever location and language (since it is possible to understand which terms represent proper names and since there are news about the location). The main focus in this paper is the application of the approach for referent disambiguation (cities with the same name) and for geo-referencing of texts where the name of the location is not present (*indirect reference ambiguity*, as defined in this work). To do that, the work utilized different corpora for identification of Location Indicators to be used in Gazetteers. The different corpora were tested and compared among them and against a baseline Gazetteer created with names of streets and neighborhoods for all Brazilian cities.

The paper is structured as follows. Section 2 discusses related work and defines the problem that is focus of this work, section 3 presents the different methods tested for automatic creation of Gazetteers (identification of Location Indicators), section 4 presents and discusses experiments and evaluations and section 5 presents concluding remarks.

2 Related Work

Some works utilize supervised learning to create Gazetteers, identifying names that are related to geographical locations from a training corpus. [Overell and Ruger 2007] utilize Wikipedia as source for identifying terms related to toponyms. The technique analyzes pages (entries) related to names of cities. The main goal is to obtain synonyms. [Popescu et al. 2008] also utilize Wikipedia to extract references to cities. [Buscaldi et al. 2006] combine Wikipedia and Wordnet as information sources; Wikipedia is useful to identify terms related to locations and Wordnet is used to identify kinds of locations and to identify in Wikipedia only the pages related to locations, eliminating ambiguities as pages related to non-geographical entities with the same name. The work of [Rattenbury et al. 2007] extracts relations between locations and terms, analyzing semantic tags registered in Flickr (http://www.flickr.com) associated to locations.

The problem is that Wikipedia, Wordnet and Flickr depend on human effort for updates. This may cause the lack of coverage (locations without information) in the Gazetteer or lack of specialized indicators (few indirect references).

The work of [Borges et al. 2003] obtain geographical information from Web pages. The technique finds indirect references as telephone numbers, zip codes and locations names in Web pages related to one city, using a tool for generation of wrappers, that has to be trained with manually selected examples.

An alternative solution is to use textual news published in the Web as source for creating and maintaining Gazetteers. The dynamic characteristic of news allows the identification of specific and up-to-date references and covering a greater number of locations.

[Ferres et al. 2004] utilize machine learning methods over news, obtaining co-referent named entities (for example, “Smith = John Smith”) and acronyms (“USA = United States of America”). [Maynard et al. 2004] utilize similar techniques over annotated corpus. [Kozareva et al. 2006] retrieve toponyms and person names using positioning expressions. They do not identify correlation between the terms and the toponyms. [Garbin and Mani 2005] utilize news to identify collocations between terms and locations. However, the window for analyzing collocations is limited to a distance of 20 terms (they do not utilize relations in the whole text). [Smith and Mann 2003] also analyze collocations in news, however they do not consider the degree of importance or weight of the relations between terms.

The problems of the cited works that utilize news include:

- the need for selecting and preparing a training corpus of news;
- the analysis of relations in windows with limited distance between terms;
- the use of relations without weight, disregarding the relative importance of the relations between terms and locations.

The contributions of the proposed work include:

- the use of news text for the training step, that is, to discover relations between terms (the discovery of Location Indicators), without the need for manually annotating a training corpus; the work does not discuss how to capture news, only suggesting the use of news texts without the need of manual annotation;
- the use of a greater window of words, considering also relations between locations and indicators present in different sentences;
- the use of a weight to determine the importance of the relations identified.

The work also evaluates the proposed approach for constructing Gazetteers in a real geo-referencing process and compares the approach with a Gazetteer created with names of streets and neighborhoods. Furthermore, the work discusses and compares different training corpus composed by news, in order to determine whether choices in the corpus selection influence the results or not.

3 The Approach for Discovering Location Indicators from News

The main goal of this work is to test an approach that identifies Location Indicators related to geographical locations, by analyzing texts of news published in the Web. The work is based on the assumption that the majority of news has some kind of Location Indicator inside the text and that statistical analysis may be utilized for retrieving news according to location data. Gazetteers are created with the identified relations and then they are utilized for geo-referencing of news. Different corpora of news are evaluated for the creation of Gazetteers (and these are evaluated by their ability to correctly identify cities in texts of news).

The first step is to collect news in the Web. In the approach, this step consider an random selection, that is, the capture the text of every news published in Web pages, without filtering. The approach does not indicate a special website or a specific technique for this selection but recommends to use websites that publish news with certainty. The suggestion is the use of well-known and reliable information sources.

The second step is the identification of relations between city names and other terms (Location Indicators). This step demands a pre-processing of the news. As “location indicators” are usually represented by proper names (PNs), the first task is to identify PNs in the texts of the news. This identification is made by analyzing words that start with uppercase, also considering special cases of multi-words (as for example, New York) and expressions that include prepositions (i.e., Massachusetts Institute of Technology). Regular expressions were defined and utilized in this task. Prepositions and adverbs that start a sentence are eliminated. Following suggestion from [Amitay et al. 2004], we obtained with statistical analysis a list of prepositions and adverbs to be eliminated. These words, that appear frequently in lowercase, are called “geographical stopwords” [Hu and Ge 2007]. A special analysis is when the name of a city is part of an expression (example: New York Mayor or, in Portuguese, Prefeito de Nova Iorque). For these cases, names of cities are extracted from the expressions by considering a list of all city names in Brazil and by analyzing the use of prepositions.

The relations between city names and location indicators are determined by a weight (numerical value, representing the importance or probability of the relation). The weight is calculated by the distance between the terms inside texts of a collection (a training corpus). The idea is to calculate the distance between the terms inside each text of the collection (local weight) and then to utilize the whole collection to determine the final (global) weight. Relations between cities are eliminated.

For the local weight calculus, the approach consider two distances: the internal distance (between terms inside the same sentence) and the external distance (between terms in different sentences of the same text). A sentence is a set of ordered terms between two final points. Formulas (1) and (2) present the calculus of the internal weight Wi_k (inside a sentence k) between a city c and a location indicator r.

$$\mathrm{Wi}_k\,(c,r) = \sum_{\substack{i=1 \\ d \le 9}}^{n} \sum_{j=1}^{m} \frac{(10 - d_{c_i r_j})}{10} \quad (1) \qquad \mathrm{Wi}_k\,(c,r) = \sum_{\substack{i=1 \\ d>9; d \le 18}}^{n} \sum_{j=1}^{m} \frac{(19 - d_{c_i r_j})}{100} \quad (2)$$

Where,

d_{xy} is the number of terms between x and y in the sentence, being that x references a city c and y references a location indicator r,

k is the k^{th} sentence in the text, where the terms appear together,

i is an index to the i^{th} appearance of the name of c in the sentence**,**

j is an index to the j^{th} appearance of the term r in the sentence**,**

n is the total number of appearances of c in the sentence,

m is the total number of appearances of r in the sentence.

For d >18, the weight ***Wi(c,r)*** is fixed to the value 0.01. The internal weight (***Wi***) must be calculated for all pairs of terms (referencing cities and location indicators) that appear together inside a sentence.

Formula (3) presents the calculus of the external weight We_t, for relations between a city c and a location indicator r present in different sentences of a text t.

$$\mathrm{We}_t(c,r) = \sum_{\substack{i=1 \\ d \le 9}}^{n} \sum_{j=1}^{m} \frac{(10 - d_{c_i r_j})}{1000} \quad (3)$$

Where,

d_{xy} is the number of sentences between x and y in the text t, being that x references a city c and y references a location indicator r,

i is an index to the i^{th} appearance of the name of c in the text t**,**

j is an index to the j^{th} appearance of the term r in the text t**,**

n is the total number of appearances of c in the text t,

m is the total number of appearances of r in the text t,

t is the text for which the external weight is being calculated.

For d > 9, the weight ***We(c,r)*** is fixed to the value 0.001. The external weight (***We***) must be calculate for all pairs of terms (referencing cities and location indicators) that appear in the text, in different sentences. The formulas and predefined values for d were defined by empirical analysis of samples of texts. The weights were established to give more relevance to closer relations (inside a sentence) but without disregarding far relations (for example, in different sentences of the text).

The local weight of a relation between ***c*** and ***r*** is calculated as the sum between the internal weight (***Wi***) and the external weight (***We***), for each text (one at each time), as exposed in formula (4). Local weight must sum all internal weights of a relation between ***c*** and ***r***, remembering that internal weights are calculated for each sentence.

$$\mathrm{Wl}_t\,(\mathrm{c,r}) = \left[\sum_{k=1}^{n} Wi_{kt}(c,r)\right] + We_t(c,r) \quad (4)$$

Where,

$\boldsymbol{Wl_t (c,r)}$ is the local weight between ***c*** and ***r*** for the ***t***th text in the collection,

t is an index for all texts in the collection,

k is an index for all sentences in the text ***t*** where ***c*** and ***r*** appear together,

n is the total number of sentences inside the text ***t*** where ***c*** and ***r*** appear together,

$\boldsymbol{Wi_k(c,r)}$ is the internal weight between ***c*** and ***r*** for the ***k***th sentence in the text ***t***,

We (c,r) is the external weight between ***c*** and ***r*** for the text ***t***.

The local weight considers relations inside each text. A global weight was defined to consider the whole collection and is calculated as exposed in formula (5).

$$\mathrm{Wg(c,r)} = \frac{\sum_{i=1}^{n} Wl_i(c,r)}{Z} \quad (5)$$

Where,

$\boldsymbol{Wl_i}$ is the local weight between ***c*** and ***r***, considering the text ***i***,

i is an index to the texts of the training collection**,**

n is the total number of texts in the training collection,

z is the total number of cities ***c*** that are related to ***r*** in the collection.

This formula normalizes the weight by dividing the sum by the total number of cities that are related to the term ***r***, considering that a term ***r*** may be related to more than one city. The argument is to give more importance for terms that are related to few cities; general relations or terms (that are related to more cities) will receive a smaller weight.

Other formulas were tested, as for example utilizing simple frequency for the relations between cities and location indicators (without weights) and not utilizing normalization (without dividing the global weight by ***z***). However, results of formal tests (previously carried out) led us to conclude that the formulas presented in this paper generates better results (for example, gains of 15% in precision).

The resulting Gazetteer to be used in posterior geo-referencing processes is composed by a set of cities, each one with a list of Location Indicators (single terms or expressions). Between the city and the indicator, there is a weight (the global weight), representing the relative importance of the relation for identifying the city when the indicator is present in the text (in the case of this paper, texts are news).

4 Experiments and Evaluations

Experiments were carried out to test the approach, including the method utilized for calculating the weight of relations between cities and Location Indicators, and also to compare different training corpus utilized for identifying these relations and thus for creating the Gazetteers.

The evaluation process consists in constructing different Gazetteers with different training corpus and then performing geo-referencing of news from a test collection captured in the Web, analyzing the ability of each Gazetteer in correctly identifying the city associate to the news, through measures like precision and recall. Each Gazetteer has the same structure: a set of cities, each one associated to a list of location indicators. Each association between a city and a Location Indicator has a weight, that is, the global weight calculated as explained in the early section of this paper.

The following Gazetteers were constructed:

(C1) 3000 NP X NP *Old:* the training corpus was composed by 3000 news published in the site Folha Online (http://www.folha.com.br), between the years 2001 and 2006; only relations between proper names were considered;

(C2) 3000 NP X NP *New:* the training corpus was composed by 3000 news published in the site Folha Online, between the years 2007 and 2008; only relations between proper names were considered; the idea is to compare this Gazetteer (with recent news) to the previous Gazetteer (with old news), but both with the same quantity of texts;

(C3) 6000 NP X NP (New+Old): the training corpus was formed by the union of both previous Gazetteers; the idea is to test if a greater collection of texts can generate better results;

(C4) 3000 NP X NP (SA): this Gazetteer was constructed from a training corpus with 3000 news recently published in Folha Online; however, the difference to the previous corpus is that this one was composed only by news that are related to only one city; the idea is to evaluate if training news with only one city result in better performance;

BASELINE: this Gazetteer was composed by location indicators corresponding to names of streets and neighborhoods of the cities. This corpus was created from a special database containing all Brazilian cities and their respective streets and neighborhoods. For this case, the global weight of the relations was not calculated and the value 1 was assumed for all relations.

For evaluating the quality of the Gazetteers (and indirectly the quality of the each corpus utilized), a collection with 230 news published in the web was utilized as a test corpus (news were randomly captured from different years from the Folha Online). No common news were utilized in training and test collections. Only 9 Brazilian cities were considered for the test, including the greatest cities and some medium cities with more than 100 thousand habitants. Each test news references only one city and has at least one location indicator. The goal is to evaluate if each Gazetteer is useful for identifying the city inside the text of a news.

Due to those restrictions (news with the presence of location indicators and published in different time period), this work utilized training and test corpora especially created for the experiments, instead of using pre-existing corpora as for example GeoCLEF[1] and HAREM[2]. The set of news utilized in the experiments are available for other authors[3].

The evaluation process consists in identifying proper names in the test texts and to compare these terms to the ones stored in the Gazetteer, remembering that it is possible that one term is associated to more than one city in the Gazetteer. Using a probabilistic reasoning, the evaluation process determines the probability of each city be present in the text. Only the more probable city is considered associated to each test text.

The probabilistic reasoning works as following:

- for each city present in the Gazetteer, the steps below are performed;
- for each term associated to the city in the Gazetteer, the presence of this term is verified in the text;
- if the term is present in the text, its weight (global weight as stored in the Gazetteer, associated to the city in question) is summed to the total probability of the city to be present in the text;
- the final sum is utilized as the probability of the city to be present in the news;
- this process is repeated for each city in the Gazetteer and for each text in the test collection;
- only the city with greater probability is considered the unique city associated to the text.

This evaluation process was done for each of the 5 Gazetteers described early.

For each text in the test collection, only one city was associated by the approach being tested. After that, the measures Precision, Recall and F1 (that combines precision and recall, with the same weight) were applied for each Gazetteer.

Results are presented in the table 1. Lines are ordered by the value of F1. The last column presents the total number of relations between a city and a term, present in each Gazetteer. Figure 1 presents the results of precision and recall in a graphical figure.

4.1 Results analysis

This sub-section analyzes the results and discusses the main points.

Comparing the four Gazetteers created by the approach against the baseline Gazetteer (created with names of streets and neighborhoods), we can note that the approach generates better results: all the four Gazetteers performed better than the

---

[1] http://ir.shef.ac.uk/geoclef/

[2] http://www.linguateca.pt/aval_conjunta/HAREM/

[3] http://gpsi.ucpel.tche.br/~cleber/geoinfo2008/

baseline in both precision and recall measures. We then conclude that news are useful for the creation of Gazetteers and also improve geo-referencing processes. News can help the identification of location indicators that are not related to streets and neighborhoods. A special analysis found that, considering the 100 location indicators with more weights for each tested city in the four Gazetteers constructed by the approach, only 19% of the terms were present in the baseline Gazetteer.

Table 1. Precision, Recall and F1 for each Gazetteer

Gazetter	Prec	Rec	F1	N. Rel.
(C3) 6000 NP X NP	100%	40%	0.5714	9159
(C2) 3000 NP X NP *new*	100%	39%	0.5612	5783
(C1) 3000 NP X NP *old*	100%	36%	0.5294	6945
(C4) 3000 (SA) NP X NP	99.3%	35%	0.5176	4757
Baseline (Streets and Neighbors)	91%	22%	0.3543	119184

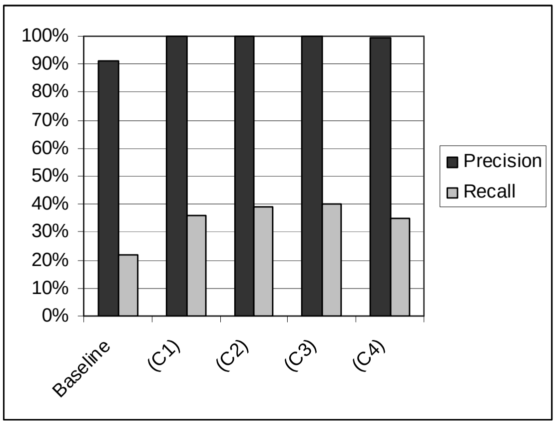


Figure 2. Graphical results of Precision and Recall for each Gazetteer

Comparing the four Gazetteers among them, we first can note that Gazetteers created with news published in different time periods (C1 vs. C2) had a small difference in performance, with a little advantage in recall measure to the Gazetteer created from more recent news (recall: C2 = 39% vs. C1 = 36%). We can conclude from this examination that more recent news are better, but we do not have to capture real-time news or even up-to-date news, because news published one year later can serve for the construction of Gazetteers with relative good performance.

Comparing training collections with different sizes (C3 vs. C1 and C2), we can note that the corpus with greater size (C3) has a better performance but with a small improvement (1.7%). This leads us to conclude that the size of the corpus is important but it may have a limit of performance. Future tests must analyze the size of the training collection composed by news.

Analyzing the performance of the Gazetteer C4, constructed from a corpus where only a city was present in each training text, we can note that this kind of corpus does not bring improvement in the performance. The initial idea was to improve recall, however this did not happen. Our explanation for this poor performance is that this kind of corpus generates a smaller number of relations than the other training collections, that is, identifying less location indicators.

5 Concluding Remarks

The main contribution of this work was to demonstrate that the construction of Gazetteers with Location Indicators instead of using names of cities, streets and neighborhoods are useful to improve geo-referencing processes. This is special important because the texts about locations frequently utilize this kind of information instead of the explicit name of the location (for example, textual news in the Web).

Furthermore, the paper demonstrated that these Location Indicators may be discovered by the analysis of news published in the Web. News can bring different Location Indicators, as for example related to very important people as mayors and authorities, related to entities as hospitals, airports, museums, universities, related to geographical places as highways, parks, constructions, buildings and so on. Most of Location Indicators are dynamic and may change along the time, and for this reason they do not appear in traditional Gazetteers.

Other contribution of the paper is that the creation of Gazetteers may be quite automatically done, by capturing news in the Web and applying the proposed approach. This may cover a great number of locations and maintain up-to-date detailed information about each location with little effort.

Furthermore, news has a special advantage that is to be more accessible than names of streets and neighborhoods. Databases with names of streets and neighborhoods are difficult to be found or must be paid. In addition, such databases, if available, may not consider new cities or changes in the existing cities (as cities that grow fast).

However, we should remember that it is necessary the existence of news about the location for the identification of Location Indicators (related terms). We believe that even small cities have newspapers or local informative vehicles (electronic or in paper) that can be used as a training collection for the Gazetteer construction.

The approach was tested with news written in Portuguese, but other languages may be utilized. The requisite is that be possible to identify proper names in the language. The rest of the approach, including the formulas, remain equal for all languages.

The paper also analyzed different corpus of news as training collections for the automatic construction of Gazetteers (evaluated by the ability of Gazetteers in correctly identifying cities in texts of news). The conclusion is that it is important to maintain the Gazetteer along the time, utilizing more recent news to update the location indicators and the corresponding weights. Although the update of the Gazetteer is important, it can be done one time per year. This is an important finding because the maintenance of the Gazetteer demands efforts and costs.

Future works include the evaluation of different sources, such as Wikipedia, scientific articles and websites for the semi-automatic construction of the Gazetteers and the evaluation of the size of the training collection.

6 Acknowledgements

This work is partially supported by CNPq and CAPES (entities of the Brazilian government for scientific and technological development).

7 References

Amitay, E., Har'el, N., Sivan, R. and Soffer, A. (2004) Web-a-where: Geotagging Web Content. In *Proceedings of the 27th SIGIR*, pages 273–280.

Borges, K. A. V., Laender, A. H. F., Medeiros, C. B., Silva, A. S. and Davis Jr., C. A. (2003) The Web as a data source for spatial databases, *V Simpósio Brasileiro de Geoinformática - GeoInfo*, Campos do Jordão (SP).

Buscaldi, D., Rosso, P. and Garcia, P. P. (2006) Inferring geographical ontologies from multiple resources for geographical information retrieval. In *Proceedings of the 3rd Workshop on Geographical Information Retrieval (GIR)*, pages 52–55, Seattle, USA.

Clough, P., Sanderson, M. and Joho, H. (2004) Extraction of semantic annotations from textual web pages. Technical report, University of Sheffield.

Delboni, T.M., Borges, K.A.V., Laender, A. H. F. and Davis Jr., C.A. (2007) Semantic Expansion of Geographic Web Queries Based on Natural Language Positioning Expressions. Transactions in *GIS*, 11(3): 377-397.

Ferres, D., Massot, M., Padro, M., Rodriguez, H. and Turmo, J. (2004) Automatic Building Gazetteers of Co-referring Named Entities. *Proceedings of the 4th International Conference on Languages Resources and Evaluation* (LREC). Lisbon, Portugal.

Garbin, E. and Mani, I. (2005) Disambiguating toponyms in news. In *Proc. Human Language Technology Conference (HLT-EMNLP)*, pages 363–370, Vancouver, BC.

Geonames. http://www.getty.edu/research/conducting_research/vocabularies/tgn/ (accessed September 02, 2008).

Hill, L. (2000) Core elements of digital gazetteers: Placenames, categories and footprints Borbinha, J. and Baker, T. (Eds.) *Research and Advanced Technology for Digital Libraries, proceedings.*

Hu, Y. and Ge, L. (2007) A Supervised Machine Learning Approach to Toponym Disambiguation. In: Scharl, A., Tochtermann, K. (eds.): *The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society*. Springer, London, 3-14.

Kozareva, Z. (2006) Bootstrapping Named Entity Recognition with Automatically Generated Gazetteer Lists, In *Proceedings of EACL student session (EACL)*, Trento, Italy.

Leidner, J. (2004) Towards a reference corpus for automatic toponym resolution evaluation. In *Workshop on Geographic Information Retrieval held at the 27th Annual International ACM SIGIR Conference (SIGIR)*, Sheffield, UK.

Leveling, J. and Hartrumpf S. (2007) University of Hagen at GeoCLEF: Exploring location indicators for geographic information retrieval. In *Results of the Cross-Language System Evaluation Campaign, Working Notes for the CLEF Workshop*. Budapest, Hungary.

Leveling, J., Hartrumpf, S. and Veiel, D. (2006) Using semantic networks for geographic information retrieval. In Peters C., Gey F. C., Gonzalo J., Jones G. J. F., Kluck M., Magnini B., Muller H., de Rijke M., editors, Accessing Multilingual Information Repositories: *6th Workshop of the Cross-Language Evaluation Forum, CLEF*, Vienna, Austria, LNCS. Springer, Berlin.

Li H., Srihari R. K., Niu C., Li W. (2003) InfoXtract location normalization: a hybrid approach to geographical references in information extraction. In *Workshop on the Analysis of Geographic References*, *NAACL-HLT*, Edmonton, Canada.

Maynard, D., Bontcheva, K. and Cunningham, H. (2004) Automatic Language-Independent Induction of Gazetteer Lists. In *Proceedings of 4th Language Resources and Evaluation Conference (LREC)*.

Overell, S. E. and Ruger, S. (2007) Geographic Co-occurrence as a Tool for GIR. In *Proceedings of the Workshop On Geographic Information Retrieval (GIR),* Lisboa, Portugal.

Popescu, A., Grefenstette, G. and Moëllic, P. A. (2008) Gazetiki: automatic creation of a geographical gazetteer. In *Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries table of contents*, Pittsburgh PA, PA, USA.

Rattenbury, T., Good, N. and Naaman M. (2007) Towards Automatic Extraction of Event and Place Semantics from Flickr Tags. In *Proc. of SIGIR*, Amsterdam, Netherlands.

Sanderson, M. and Kohler, J. (2004) Analyzing geographic queries. In *Proceedings of the Workshop on Geographic Information Retrieval*, Sheffield, UK.

Smith, D. and Mann, G. (2003) Bootstrapping toponym classifiers. In *Workshop on the Analysis of Geographic References*, *NAACL-HLT*, Edmonton, Canada.

TGN (Getty Thesaurus of Geographic Names). http://www.getty.edu/research/conducting_research/vocabularies/tgn/ (accessed September 02, 2008).

Global Geoportal for Remote Sensing Images Centers

Vanessa Cristina Oliveira de Souza, Gilberto Câmara

National Institute For Space Research
Av dos Astronautas, 1.758 - Jd. Granja - CEP: 12227-010
São José dos Campos – SP - Brazil

vanessa@dpi.inpe.br, gilberto.camara@inpe.br

Abstract. *The objective of this work is to propose a geoportal capable of integrating the catalogues of the different images centers without these centers having to migrate their current computational architectures to a specific architecture. In this work, the mediator architecture is considered a flexible and efficient way of building a global geoportal for remote sensing images. The three components of this architecture, portal, mediator and data sources, will be analyzed and specified. The web services technology, as a solution for implementing this architecture, will also be analyzed. As a demonstration prototype, we have integrated the CBERS data distributed by INPE in an existing geoportal.*

1. Introduction

Remote sensing images are used in various fields with various objectives. During the last decades of the 20th century, the repositories of remote sensing images, the images centers, organized their data in offline tape archives and each image was individually generated for the user. At the end of the 1990's the centers changed the way their images were distributed. With the popularization of the Internet, these repositories began to convert their data to online access.

Each center created its own images catalogue using its own interfaces, software and hardware platforms. It became difficult for the user to interact with different interfaces and to find the catalogues online since the current search engines were not designed to find geographic information. Therefore, the user must find the catalogue, learn how to use it and later combine the data manually.

In this context, a geoportal where the various images centers around the world could be accessed, making the particularities of each one clear to the user, is desirable.

In this work we will analyze the problem of organizing *online* distribution so that the centers can work in cooperation, increasing the amount of information available to the user. We will try to answer the question: How can we conceive and build a geoportal with search and recovery services for remote sensing images centers that function in cooperation?

We assume that a mediated architecture [1] is a flexible and efficient way of building the geoportal for the images centers. We will study the concept and building of mediators based on the *web services* technology [2-5] as a solution to the problem.

After studying some existing online images catalogues [6] and analyzing the necessary components of this mediated architecture, we propose a solution based on the current *web services* technologies. We also carried out an experiment that integrated the data from the CBERS[1] images catalogue, distributed by INPE[2], in an existing geoportal (eoPortal) to validate the use of the mediated architecture and to analyze the positive aspects and the deficiencies of this geoportal.

2. Theoretical References

A geo-spatial portal (geoportal) is the human interface for a collection of *online* geo-spatial information, including services and datasets. Technically speaking, geoportals are *sites* connected to *web* servers that provide metadata on their geographic data or services [7, 8]. For Tait [9], geoportals provide applications that can search, map, publish and administrate geographic information.

Data integration is the problem of combining data from different sources, offering the user a unified vision of them (global schema), and defining a set of queries that determine how each element of the global schema is obtained in function of the data stored in the local data sources [10, 11]. A schema for the mediated is an unified vision available from the mediator. Queries can be made against that schema and submitted to the mediator. So, they are decomposed at run time into queries on the local data sources. The results from these queries on the local data sources are translated, filtered and merged, and then the final answer is returned either to the user or to the application[12].

The data integration systems follow two main approaches: the materialized approach and the virtual approach. The virtual approach is the most applied to our problem, as the data remain in the sources and the information is extracted directly from them when a consult is requested. The main advantages of the virtual approach are the non-replication of the data and the fact that the information recovered is always updated. The disadvantages of this approach include a possible inaccessibility of the sources and the long response time [12-14].

The virtual approach is generally modeled using the mediated architecture [1]. In this architecture there is a *software* module, the mediator, that receives and treats the searches submitted to the integration system, decomposing these queries into sub-queries that will be submitted to the data sources [15]. The *wrapper* is a program used to make the translation between the sources' data model and the data model defined by the mediator (Figure 1).

The Web Services (WS) are currently the most promising way to integrate applications on the Internet [3]. A web service is a software application identified by a URI, whose interfaces and binding are capable of being defined, described and discovered by XML artifacts and supports direct interactions with other software applications using XML based messages via internet-based protocols [16]. The virtual approach modeled by mediated architecture is generally implemented by web services.

---

[1] China-Brazil Earth Resources Satellite - http://www.cbers.inpe.br/
[2] Brazilian National Institute For Space Research

In W3C WS, XML is used to encode data. The protocol of communication is SOAP (Simple Object Access Protocol). The services' description is standardized by the WSDL (Web Services Description Language) and the services' discovery by UDDI (Universal Description, Discovery and Integration).

Compared with other protocols, the SOAP has a low performance, considering that the messages are described in text (XML), while in RPC (Remote Procedure Call) systems, they are exchanged in binary format. Govindaraju et al. [17] say that the SOAP message is four to ten times larger than the binary representation the same.

The eoPortal[3] is the main entrance of an architecture made available by ESA[4] called Service Support Environment (SSE), in which the users and the service providers can interact automatically and dynamically. In this portal, the providers register and offer their services and data and the users access them [18].

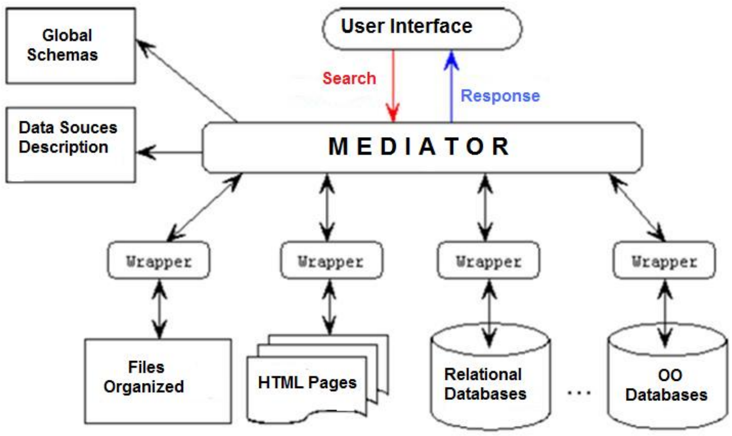


Figure 1 : Virtual approach modelled by mediators.

The SSE is a system open to remote integration of services with domains focused on observation of the Earth and on the geographic information systems. The architecture of the SSE can be seen in Figure 2. In the system, services can be integrated without having to re-compile or re-design it [19]. Its architecture is constituted by a set of functionalities, components and gateways, providing a distributed environment oriented towards services. This architecture is non-proprietorial, open, scaleable and robust for discovering and distributing services.

In the area aimed at service providers, presented in Figure 2, it is the Toolbox, a development environment provided by SSE to help the providers to create their services.

3. Global Geoportal for Remote Sensing Images Centers

We will begin this geoportal proposal with its architecture. We propose a mediated integration architecture based on the virtual approach and implemented by web services. Therefore, the mediator component in Figure 1 is a web service, as are the wrappers. The sources are described by WSDL and UDDI. The integrated vision is a schema

[3] http://eoportal.gov/
[4] European Spacial Agency

XML that defines the content of the SOAP messages exchanged between the mediator and the wrappers (Figure 3).

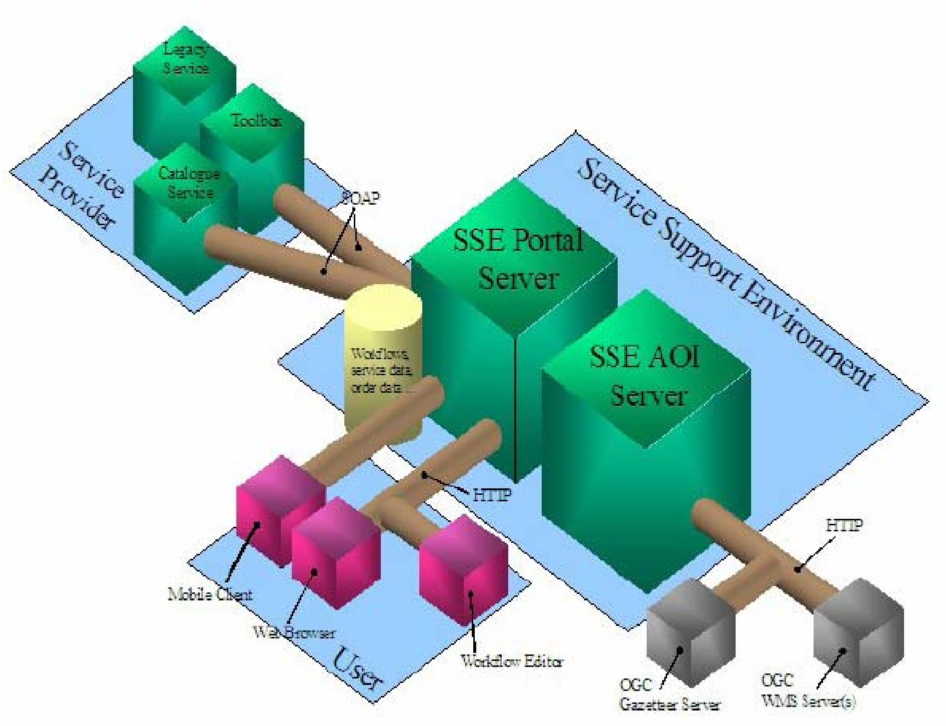


Figura 2 : SSE architecture.

The architecture is composed of: portal, mediator and data sources. The mediator component of this architecture is detailed in Figure 4. As the SOAP, the virtual approach has performance problems. Therefore, the purpose of the mediator filter is to optimize the transmission of the messages, sending them only to the sources that can actually respond to the consult made. This filtering is done using the WSDL and UDDI. Ideally, the sources would register the image sensors, the area of coverage of their scenes and the period of time in the UDDI. After filtering the sources, the messages are sent to them and the results are integrated into one result, which is sent back to the portal.

The sources have their own data models and to transform them into the integrated vision, and to standardize the SOAP messages exchanged between the mediator and the sources, a web service wrapper must be implemented. Basically, the operations of this service follow Figure 5. After receiving an xml message with the search parameters, the service must decode this xml to remove the parameters, consult the database and write a new xml message, which will return the result of the consult to the geoportal.

All the interaction with the user occurs in the portal component of Figure 3. It is a web interface that functions as a client of the mediator component. This portal contains:

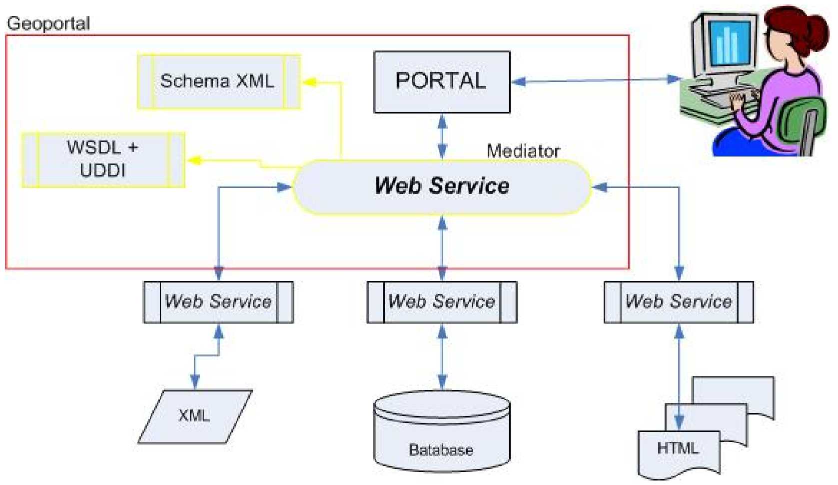


Figure 3: Architecture proposed for the global geoportal implemented by web services.

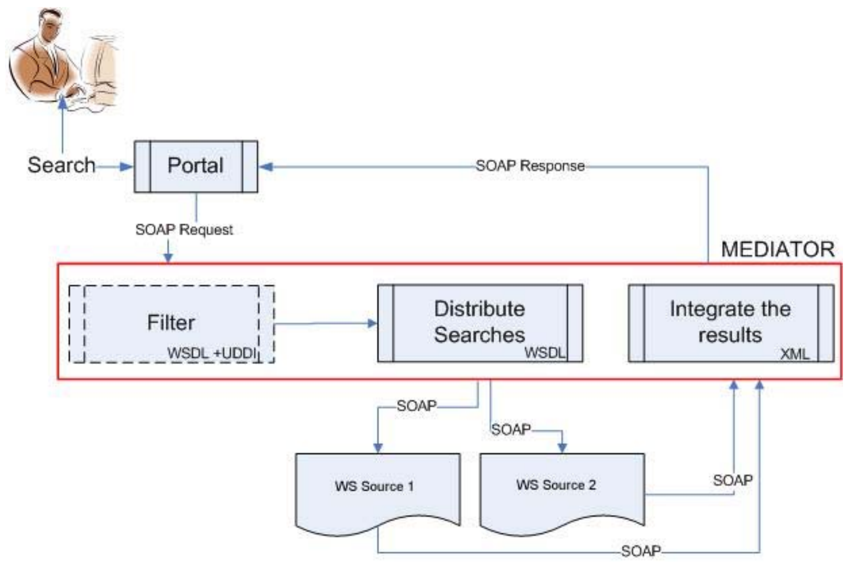


Figure 4 : Mediator component.

a) Images catalogue:

Catalogues publish metadata and provide mechanisms to consult and recover information from distributed repositories. This component is a web interface which composes queries and presents the information from the repositories to the user, via a graphic interface. We can, however, associate a data access interface to the catalogue, from where the user could obtain the actual scene.

The metadata presented in this work are the result of a study of major catalogues of images available today [6]. The ISO 19115 [20] was also used in this work. In the tables below, the metadata were divided into mandatory, optional and additional fields. This division was made to evaluate the catalogues. At the interfaces of this proposed catalogue, all metadata should be presented.

The search interface must contain the maximum number of methods for locating the area of interest, such as gazetter, geo-coding, editing map, file upload and coordinate entry. The main role of this interface is to help the user find the data he is searching for in the best possible way. We suggest that all the parameters in Table 1 should be included in this interface. These fields will be the parameters used by the data sources during a consult.

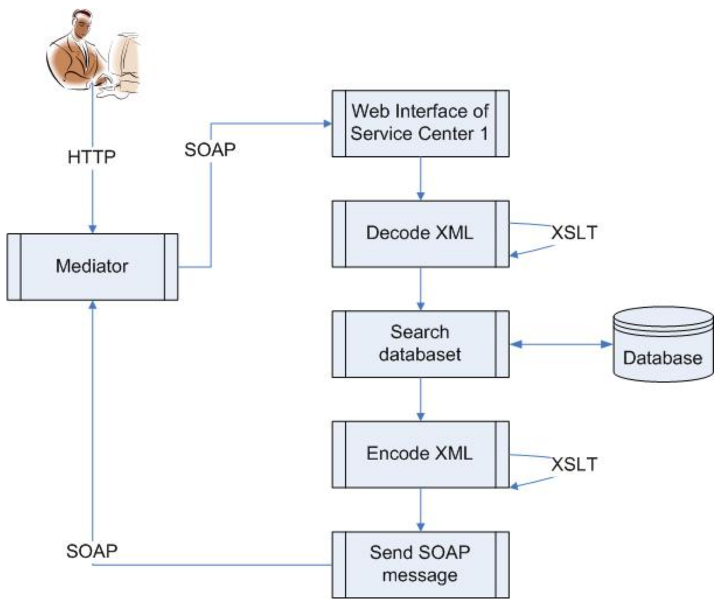


Figure 5 : Images centers' services.

Apart from these parameters, common in existing images catalogues, we also propose an interface that uses parameters that are less punctual and more semantic. Therefore, instead of searching for the sensors by their names, the user would search for them according to spatial, temporal, radiometric and spectral resolution, as well as according to a radar's polarimetry. This form of consult does not require knowledge of the sensor system, allowing the user to search for any sensor that best meets his needs, making the interface more flexible and intelligent.

The interface for publishing the metadata will present to the users the metadata returned by the consult. We classify these metadata in Table 2. The fields referenced with * were taken from ISO 19115 [20].

Table 1: Mandatory, optional and additional fields of a geoportal's search interface for RS images centers.

Mandatory Fields	Optional Fields	Additional Fields
Geographic Location	Sensors	Illumination and Azimuth angles
	Initial and final dates	Quality of the image
	Cloud coverage	Source of the data
	Online availability of the product	Path/row
		SceneId

Apart from these parameters, common in existing images catalogues, we also propose an interface that uses parameters that are less punctual and more semantic. Therefore, instead of searching for the sensors by their names, the user would search for them according to spatial, temporal, radiometric and spectral resolution, as well as according to a radar's polarimetry. This form of consult does not require knowledge of the sensor system, allowing the user to search for any sensor that best meets his needs, making the interface more flexible and intelligent.

The interface for publishing the metadata will present to the users the metadata returned by the consult. We classify these metadata as in Table 2. The fields referenced with * were taken from ISO 19115 [20].

Table 2 : Mandatory, optional and additional fields of a geoportal's interface for presenting results for RS images centers.

Mandatory Metadata	Optional Metadata	Additional Metadata
Scene Identifier	Format of the image	Condition of the image*
Geographic coordinates	Cost	Quality of the image*
Path/Row	Source of the data	Level of processing*
Day and time of the acquisition	Órbit	Radiometric calibration data *
Sensor system	Nadir	Camera calibration data *
Cartographic Projection	Direction of the orbit	Distortion data*
Form of access to the data	Azimute *	Triangulation*
Cloud coverage	Elevation*	

The interface proposed presents all the metadata in Table 2, divided in categories. All the metadata have a key that explicits their function and possible values, assisting less experienced users. All the scenes must be drawn over a reference map. The mandatory metadata must be presented for all the scenes resulting from the consult. The others must be presented only if requested by the user. The order in which the scenes appear must be decided by the user. By default, the more recent scenes appear first.

Coupled with the interface for publishing the metadata there must be an interface that allows the user to request the desired scene. This interface should send the requests to the provider centers and deliver the products to the user of the geoportal. The scenes may be downloaded or delivered to the user's e-mail.

b) Administrative interface:

The administrative interface will manage the registration of the users and of the images centers. The registration of the users will enable them to make full use of the geoportal's functions, as well as composing the portal's statistics. The registration of the centers is also important since they then gain the status of service providers and can publicize their data.

c) Extra functions:

Many functions can be added to the geoportal to attract and help users. We suggest the list below. Only the last can be used by users not registered in the geoportal:

- **Storing search parameters and consult results:** this function is present in the Earth Explorer5 and its objective is to save time. Saving the search criteria is interesting for users who always make the same consult, changing only the date, for example.
- **Forum:** as it is a niche environment, it is common for the portals to have forums so that theirs users can communicate with each other, exchange information and clarify doubts.
- **News:** the objective of this function is to keep the users updated on the news of the geoportal itself or other news in the field of RS images. In this way, the user can be informed, via e-mail, of a new content or a new data source in the portal.
- **Compare sensors:** this function helps the user to choose one or another sensor based on the comparison of some characteristics, such as, for example, the sensors' spatial resolution and the cost of the product.

4. Integration of the CBERS catalogue in the eoPortal

4.1. Current architecture of the CBERS catalogue

The cooperation between the Brazilian and Chinese governments established in 1988 led to the development of imaging satellites, which help monitor the natural resources of both countries. In Brazil, the CBERS program includes a policy of free distribution of the images, first inside the national territory and, more recently, to other South-American countries. This distribution is done through a catalogue available on the Internet.

The current catalogue is built on a structure based on PHP scripts and mysql relational database. The catalogue's search interface does not have an editing map, a geo-coding system and file upload. Although the catalogue provides data from many sensor systems, the searches are not multi-sensor. The interface for publishing the metadata lacks a reference map and the obligatory metadata available is minimal.

In the data access interface all communication with the user is done through e-mails. When the scene requested is ready, it is copied to an area on the Internet. The area's address is sent to the user, who has five days to complete the download.

4.2. Data model for the unified vision

The Interface Control Document (ICD) defines the external interfaces of the SSE, particularly the interfaces between the SSE Portal and the remote services of the providers. For this work, the most relevant interfaces are ESRIN EOLI [21] and MASS. EOLI defines the messages that are exchanged between the SSE and the catalogue image services that implement the "Search" and "Present" operations. MASS defines many operations, among them the operation for carrying out orders through the portal, "Order".

Communication between the portal and the catalogue service is established using SOAP on HTTP. The SSE provides pre-defined workflows that make the

[5] http://edcsns17.cr.usgs.gov/EarthExplorer/

conversion between the XML messages of the SSE Portal and the EOLI interface messages [19, 21].

4.3. Integration

The CBERS catalogue was integrated in the eoPortal through the implementation of a web service, whose operations are: Search, Present and Order. This service will perform the function of wrapper between the CBERS relational data model and the interfaces EOLI and MASS (Figure 6).

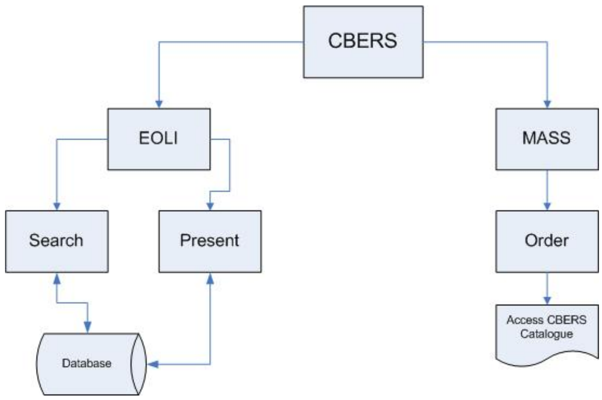


Figure 6: Implemented CBERS service.

We used the Toolbox to implement the service, which runs on an Apache Tomcat server. The operations implemented are all synchronous and are detailed below.

The scenes' metadata recovered from the database are translated to XML, following EOLI, by a second XSL script. This script is the actual wrapper, as it literally translates the CBERS catalogue schema into the EOLI data model. The XML generated is encapsulated in a SOAP message by Tomcat and sent back to SSE, which applies a style to it, codifies the scenes coordinates in GML[6] and presents both the textual and the graphic data in the eoPortal.

The Present operation functions in the same way as the Search. However, in the Present operation, the search parameter will always be the identifier of one single scene and the response requested to the database contains a larger number of metadata.

The operation Order follows the MASS interface data model. It receives as entrance parameter the identifiers of the desired scenes, the user's name and e-mail, and returns the scene, through a link, to the file stored in disk. This operation has been implemented, but not yet operationally. The service (Figure 7) can be accessed in: http://services-test.eoportal.org/portal/service/ShowServiceInfo.do?serviceId=DF80CA80.

At the moment, the service is in the test phase. The liberation for all community will be done in December.

[6] Geography Markup Language

5. Conclusions

This work discussed the potential benefits of integrating the different remote sensing images centers. We indicated how this integration can be achieved using a mediated architecture and we made a demonstration prototype.

Analyzing the current major catalogues, we observed how the integration was necessary. As the catalogues already exist in their software and hardware platforms, it is important to have a mediated architecture that unites them in one simplified interface. This work has proved that it is possible to build a mediated architecture that functions well, since the aspects that are common to all the catalogues are stronger than the differences between them. The integration between the eoPortal and the CBERS catalogue is an example of this.

The SSE is an open, free system based on free technologies that provides a high quality support. Providing services using the system does not imply any costs. In the case of images catalogues, the great advantage is the EOLI interface.

The greatest difficulty to integration with eoPortal lies in the quantity of languages that must be mastered (TSCRIPT, XML, XSL, XSLT, Xpath, WSDL, UDDI) and in translating local metadata into EOLI. There are no notation or semantic difficulties, but there is a difficulty in storing the information in the same standard as EOLI.

For the users, the implementation of this geoportal would facilitate the search for and recovery of the images and would increase the volume of data available. For the centers, this geoportal would decrease the effort of implementing and maintaining geographic web interfaces, as the providers would need only to keep the database updated and implement the wrapper service. It is essential to use the XML, and all the technology associated to it, for this architecture to function with total interaction between the parts.

Figure 7 : INPE Catalogue Service in eoPortal.

6. References

1. Wiederhold, G., *Mediators in the Architecture of Future Information Systems* IEEE Computer, 1992. **25** (3): p. 38-49.

2. Breitman, K., M. Casanova, and W. Truszkowski, *Semantic Web – Concepts, Techonologies and Applications*. 2007, London: Springer. 327.

3. Curbera , F., et al. (2002) *Unraveling the Web Services Web - An Introduction to SOAP, WSDL, and UDDI*. IEEE Internet Computing **Volume**, 8

4. Newcomer, E., *Understanding Web Services - XML, WSDL, SOAP and UDDI*. 2002, Boston: Addison-Wesley. 332.

5. W3C, W.W.W.C. *Web Services Architecture*. 2004 [cited 2007 Janeiro/2007]; Available from: http://www.w3.org/TR/ws-arch/

6. Souza, V.C.O., *Geoportal Global para Centros de Imagens de Sensoriamento Remoto*, in *Sensoriamento Remoto*. 2008, INPE: São José dos Campos/SP. p. 99.

7. Maguire, D.J. and P.A. Longley, *The emergence of geoportals and their role in spatial data infrastructures*. Computers, Environment and Urban Systems, 2005. **29**(1): p. 3-14.

8. OGC, O.G.C. *Geospatial Portal Reference Architecture*. 2004 [cited 2007 Janeiro/2007]; Available from: http://portal.opengeospatial.org/files/?artifact_id=6669.

9. Tait, M.G., *Implementing geoportals: applications of distributed GIS.* Computers, Environment and Urban Systems, 2005. **29**(1): p. 33-47.

10. Lenzerini, M., *Data integration: a theoretical perspective*, in *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. 2002, ACM: Madison, Wisconsin.

11. Lóscio, B.F., A.C. Salgado, and V.M.P. Vidal. *Using Agents for Generation and Maintenance of Mediators in a Data Integration System on the Web*. in *XVI Simpósio Brasileiro de Banco de Dados*. 2001. Rio de Janeiro - Brazil.

12. Lóscio, B.F., *Managing the Evolution of XML-based Mediation Queries*, in *Doutorado em Ciências da Computação*. 2003, Universidade Federal de Pernambuco: Recife. p. 159.

13. Barbosa, A.C.P., *Middleware para Integração de Dados Heterogêneos Baseado em Composição de Frameworks*, in *Doutorado em Informática*. 2001, PUC-RIO: Rio de Janeiro. p. 167.

14. Batista, M.d.C.M., *Otimização de Acesso em um Sistema de Integração de Dados através do Uso de Caching e Materialização de Dados*, in *Mestrado em Ciências da Computação*. 2003, Universidade Federal de Pernambuco: Recife/PE. p. 128.

15. Moura, S.L.d., *Uma arquitetura para Integração de Repositórios de Objetos de Aprendizagem baseada em Mediadores e Serviços Web*, in *Mestrado em Informática*. 2005, PUC-Rio: Rio de Janeiro. p. 158.

16. W3C, W.W.W.C. *Web Services Architecture Requirements*. Web Services Architecture Working Group 2002 [cited 2007 12/03/2007]; Available from: http://www.w3.org/2002/ws/arch/2/wd-wsawg-reqs-03262002.xml.

17. Govindaraju, M., et al. *Toward Characterizing the Performance of SOAP Toolkits*. in *5th IEEE/ACM International Workshop on Grid Computing (GRID'04)*. 2004. Pittsburgh, USA: IEEE Computer Society.

18. SSE Team. *SSE Infocenter*. 2007 [cited 2007; Available from: https://services.eoportal.org/infocenter/index.jsp.

19. SSE Team. *SSE Interface Control Document*. 2006 [cited 2007; 156]. Available from: http://services.eoportal.org/massRef/documentation/icd.pdf.

20. ISO, *Geographic Information - Metadata.* 2003, International Organization for Standardization. p. 140.

21. Nencioni, P. *Earthnet Online XML Front-End*. 2006 [cited 2007; Available from: http://earth.esa.int/rtd/Documents/EOLI-XML-ICD.pdf.

SIGWeb Builder: Uma Ferramenta Visual para Desenvolvimento de SIG Webs

Helder Aragão, Jorge Campos

GANGES – Grupo de Aplicações e Análises Geoespaciais

Mestrado em Sistemas e Computação - UNIFACS – Salvador – BA - Brasil

helderaragao@yahoo.com.br, jorge@unifacs.br

Abstract. *This work aims at the implementation of a visual tool to support the development of Web Geographic Information Systems (WebGIS) based on freeware environments. The proposed tool will minimize the effort of developing such applications and allow developers to concentrate their efforts in the functionalities and layout of the application and abstract away from the details of the target environment used to support the application.*

Resumo*. Este trabalho tem como objetivo a criação de uma ferramenta visual para o desenvolvimento de Sistema de Informações Geográficas para Web (SIG Web) baseada em ambientes livres. A ferramenta proposta irá minimizar o esforço do desenvolvimento deste tipo de aplicação e permitirá que os desenvolvedores concentrem seus esforços na definição das funcionalidades e aparência da aplicação e abstraiam os detalhes do ambiente alvo utilizado para suportar a aplicação.*

1. Introdução

Os Sistemas de Informações Geográficas para a Web (SIG Web) têm experimentado um formidável crescimento nos últimos anos, tornando-se um dos recursos mais utilizados na disponibilização e disseminação de dados geográficos na Internet. Os SIG Webs, em geral, são caracterizados por uma interface fácil e intuitiva para a apresentação de mapas temáticos. A esta característica são acrescentadas algumas funcionalidades básicas para manipulação e controle do conteúdo apresentado no mapa. A facilidade de uso pelo público não especializado em conjunto com o poder da comunicação cartográfica, constituem-se como os principais fatores na popularização desses sistemas.

Embora os SIG Webs possuam uma grande aceitação do público em geral, o seu desenvolvimento não é uma tarefa trivial. Detalhes como a tecnologia Web embutida na aplicação e as funcionalidades a serem ofertadas são fatores decisivos na escolha do ambiente de base do SIG Web. Associado a estas questões, existe atualmente uma proliferação de ambientes, proprietários e livres, para o desenvolvimento desses sistemas. As soluções proprietárias possuem recursos mais amigáveis e facilitam o desenvolvimento. Estes recursos, entretanto, fazem parte geralmente de um pacote de soluções muito mais abrangentes e com custos proibitivos para a maioria das organizações. As soluções em ambientes livres, por outro lado, não apresentam custos para a sua utilização, mas exigem que o desenvolvedor possua mais do que conhecimentos básicos em computação e domine profundamente os detalhes da arquitetura, dos arquivos de configuração e da sintaxe dos recursos disponíveis no ambiente. Esses fatores geram custos adicionais no processo de desenvolvimento e afastam as pessoas responsáveis pela produção do conteúdo da tarefa de produzir o sistema.

Independente do ambiente utilizado nos SIG Webs, o desenvolvimento de sistemas com alguma complexidade tende a se tornar um processo repetitivo e trabalhoso. Considerando que o desenvolvedor domine todos os recursos do ambiente, a sintaxe e a semântica dos marcadores e as tecnologias Web utilizadas, o processo de criação através da edição do código fonte é um processo bastante rudimentar e tedioso. Erros na codificação e configuração do ambiente, replicação de código e dificuldade na manutenção e evolução dos sistemas são algumas características inerentes a este processo rudimentar de desenvolvimento de sistemas.

Este trabalho propõe o desenvolvimento de uma ferramenta visual que apóie a construção de SIG Webs baseados em ambientes livres. A ferramenta proposta visa permitir que os usuários possam abstrair dos detalhes da implementação dos sistemas e possam, através de uma interface gráfica e intuitiva, construir SIG Webs baseado em ambientes livres. O restante deste artigo está estruturado da seguinte forma: na seção 2, são apresentadas as principais características de dois ambientes livres para publicação de mapas na web suportados pela ferramenta. A seção 3 discute os principais recursos disponíveis na ferramenta. Na seção 4, são apresentados os trabalhos futuros.

2. Ambientes Livres para Publicação de Mapas na Web

Existem atualmente algumas soluções baseadas em ambientes livres para a produção de SIG Webs. Os SIG Webs baseados nesses ambientes suportam diferentes níveis de interatividade e funcionalidade [Peng e Tsou 2003]. Estas diferenças estão diretamente ligadas à arquitetura e as tecnologias utilizadas por cada ambiente. Dentre os ambientes livres disponíveis no momento, dois merecem destaque: MapServer [MapServer 2007] e AlovMap [AlovMap 2007]. O MapServer é um dos servidores de mapas mais utilizados atualmente, possui código aberto e utiliza os recursos básicos dos navegadores para a Internet. O AlovMap é um ambiente livre baseado em aplicações Java (*applet* e *servlet*) e permite um alto grau de customização através de tecnologia XML (*Extensible Markup Language*). As principais características desses ambientes são detalhadas a seguir.

2.1 MapServer

O MapServer é um ambiente de desenvolvimento para a disponibilização de dados espaciais em ambiente Web. Esse ambiente atualmente é mantido por um grupo de desenvolvedores e pertence ao Open Source Geospatial Foundation (OSGeo) [MapServer 2007]. O MapServer suporta diversos tipos de dados vetoriais, como *Shape* da ESRI e PostGIS. Dois formatos Raster também podem ser manipulados nativamente: Geotiff e eppl7. O MapServer é compatível com as especificações do *Open Geospatial Consortium* (OGC) e suporta os serviços Web Map Server (WMS), Web Feature Server (WFS) (somente para consulta) e Web Coverage Server (WCS) [Kropla 2005; Tu 2006].

Um SIG Web baseado em MapServer possui três componentes básicos: *MapFile*, *Template* e uma aplicação *CGI* (Common Gateway Interface) [Kropla 2005]. O *MapFile* é um arquivo texto que define as características do projeto, tais como, o tamanho do mapa e as camadas de informação (*layers*) disponíveis. O *Template* é um arquivo HTML que define a organização dos mapas e dos componentes funcionais do sistema. O programa *CGI* é responsável por ler e processar os arquivos *MapFile* e *Template*.

Uma aplicação CGI é uma forma de disponibilização de conteúdo dinâmico para a Web. Quando o usuário faz uma solicitação ao navegador, este a repassa ao servidor, que identifica a requisição como pertencendo a um programa CGI. Uma vez identificada a requisição, o servidor a repassa a um programa externo, chamado de aplicativo CGI que fica responsável por prover funcionalidade requisitada pelo cliente [Miranda 2003; Peng e Tsou 2003].

A figura 1 mostra um exemplo de uma aplicação desenvolvida em MapServer. Nesta figura destacam-se quatro regiões: área de painel de controle, com o modo de operação do servidor (navegação ou consulta) e os comandos de interação com o mapa (*zoom-in*, *zoom-out* e pan), área de *layers*, área com legenda e escala e, por fim, a área do mapa.

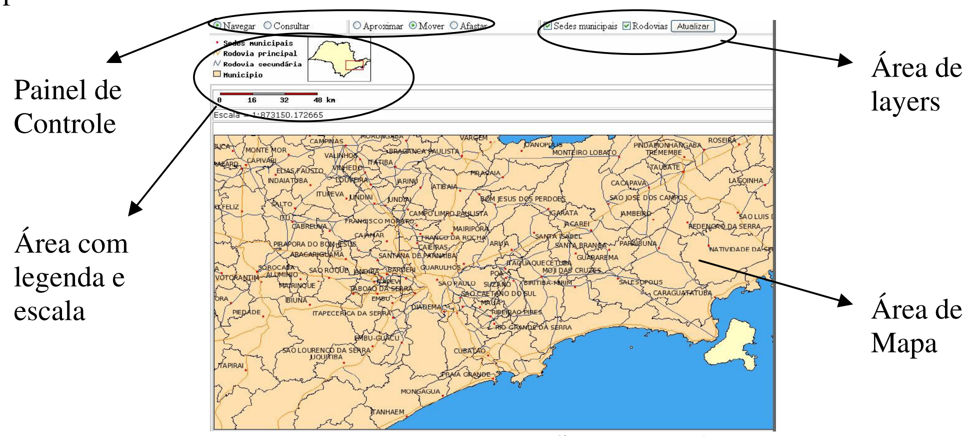


Figura 1 – Exemplo de aplicação com MapServer.

2.2 AlovMap

O AlovMap é um ambiente portável e gratuito, desenvolvido pela Universidade de Sydney, que permite a publicação de mapas na Web [AlovMap 2007; Cecarreli 2006]. Esse ambiente suporta também formatos raster e vetoriais e pode funcionar como *applet* ou *servlet* [AlovMap 2007; Babu 2003].

Uma aplicação SIG Web baseada em AlovMap possui dois componentes básicos: Arquivo de Projeto e Arquivo de Layout. Estes arquivos são codificados na linguagem XML. O arquivo de projeto define todas as características do SIG Web, tais como: *layers*, mapas temáticos, tamanho e cor dos símbolos dos *layers*. O arquivo de *layout* define a organização dos componentes gráficos da aplicação [Babu 2003].

As principais funcionalidades de um SIG Web baseado em AlovMap são: criação de mapas temáticos, visualização de *layers* e execução de *zoom-in*, *zoom-out*, *pan* e *seleção*. Uma aplicação típica com AlovMap possui quatro regiões: painel de controle, quadro com os *layers*, o mapa e o painel de status. No painel de controle ficam os botões e ferramentas para interação com o mapa, como *zoom-in*, *zoom-out*, *pan*, seleção e a lista de mapas temáticos. O Quadro de *layers* mostra os *layers* e a legenda dos mapas temáticos. No Quadro de Mapa, pode ser visualizado tanto o mapa, como os temáticos. O Painel de Status possui um botão de LOG, onde são registrados possíveis erros. Neste último painel, são exibidas mensagens sobre o *layer* ativo, coordenadas e a quantidade de zoom que está sendo utilizado. A figura 2 mostra um exemplo de uma aplicação típica utilizando o AlovMap [Aragão, Britto, et al 2007].

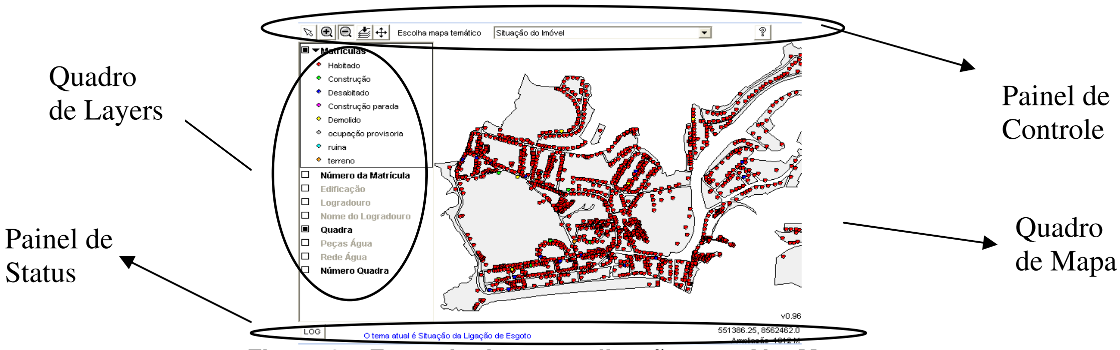


Figura 2 – Exemplo de uma aplicação com AlovMap.

3. Proposta da Ferramenta

A principal característica de um SIG Web é disponibilizar informações geográficas de uma forma gráfica e intuitiva. Desta maneira, mesmo os usuários que não possuem treinamento especializado, conseguem absorver informações que, de outra forma, só eram absorvidas por técnicos especializados com domínio na utilização de alguma ferramenta SIG proprietária.

De forma a possibilitar e incentivar o consumo de informações geográficas, as funcionalidades ofertadas nos SIG Webs devem se restringir às tarefas básicas, intuitivas e de fácil assimilação por usuários não treinados. Consideramos como funcionalidade mínima a ser oferecida nesses sistemas a escolha das informações, mostradas no mapa, as funções de manipulação como *zoom* e *pan* e pelo menos uma função de seleção que permita obter informação detalhada de uma determinada feição no mapa.

Mesmo disponibilizando um conjunto restrito de funcionalidades, o desenvolvimento de um SIG Web ainda é uma tarefa trabalhosa, enfadonha e requer conhecimentos específicos que não estão diretamente ligados ao domínio SIG, mas sim a área de Computação. Visando minimizar o esforço de desenvolvimento de um SIG Web e propiciar que pessoas sem grandes conhecimentos das tecnologias para aplicações na Internet possam criar o seu próprio sistema, este trabalho propõe o desenvolvimento da ferramenta SIGWeb Builder, uma ferramenta visual para o desenvolvimento de SIG Webs.

O SIGWeb Builder é uma ferramenta visual que permite a definição da estrutura de um SIG Web baseada em componentes gráficos. Com esta ferramenta, o desenvolvedor SIG Web deve se preocupar, basicamente, com a apresentação do sistema, com as funcionalidades SIG que serão disponibilizadas e com a definição do ambiente alvo que será utilizado como suporte para publicação dos mapas. A ferramenta SIGWeb Builder gera automaticamente e de forma transparente para o desenvolvedor todo o código para o ambiente alvo selecionado.

Foram definidos, inicialmente, dois ambientes alvos para a publicação do projeto: MapServer e AlovMap. A escolha destes ambientes foi motivada pela grande aceitação dos mesmos, por serem ambientes livres, por possuírem arquivos de configuração com sintaxe bem definida e por serem baseados em arquitetura Web distintas. Enquanto o MapServer permite a publicação utilizando CGI, o AlovMap utiliza *Applets* e *Servlets* para a disponibilização de mapas em ambiente Web.

Neste estágio, o SIGWeb Builder possui um conjunto de componentes para definição do projeto e uma área para visualização prévia da aparência do SIG Web (figura 3). Os componentes para definição do projeto são agrupados de acordo com suas funcionalidades em cinco paletas: Projeto, Layers, Navegação, Mapa e Utilidades.

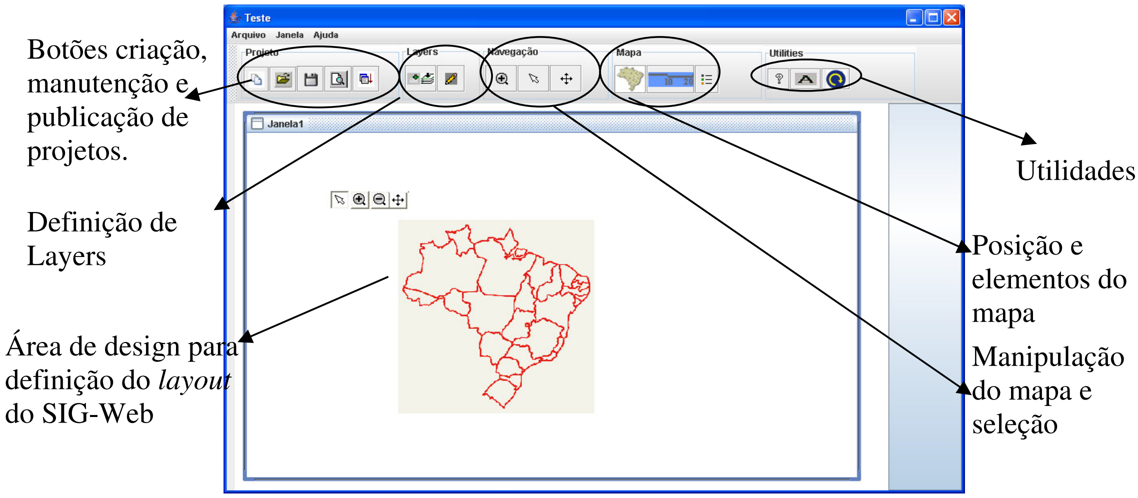
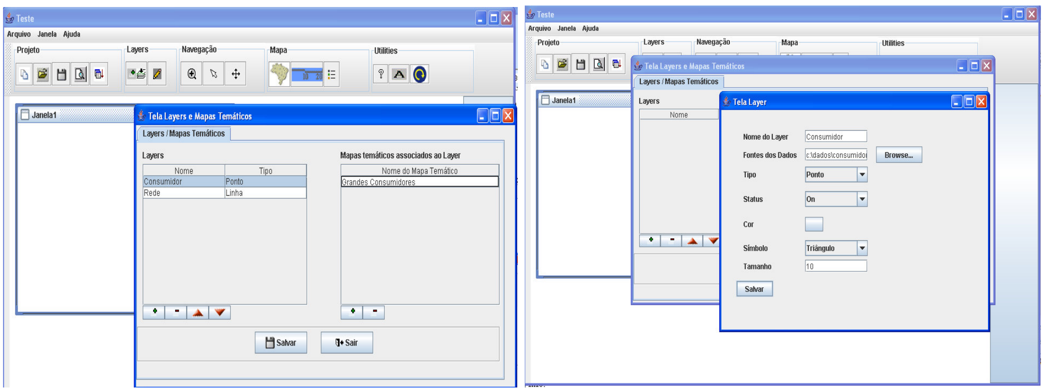


Figura 3 – Tela Principal do Projeto

A paleta de Projeto permite a criação de um novo projeto ou manutenção de projetos existentes. Nesta paleta, é possível também publicar o projeto, isto é, gerar os arquivos necessários a execução do SIG Web. Para esta tarefa, entretanto, é necessário definir o ambiente alvo desejado.

A paleta *Layers* permite definir quais os *layers* e os seus mapas temáticos a serem disponibilizados (figura 4.a). A aplicação permite definir a ordem, a fonte de dados e tipo do *layer* (ponto, linha ou polígono) (figura 4.b). É possível também definir as expressões e cores utilizadas para a geração dos mapas temáticos.



a) b)

Figura 4 – Telas de configuração de layers: a) definição de *layers* e mapas temáticos e b) definição das características do *layer*

A paleta de Navegação (figura 3) permite a inclusão no SIG Web dos elementos gráficos que ativam as funcionalidades de zoom (in/out), pan e o modo de seleção. Esta última permite que o usuário obtenha informações detalhadas de um ponto selecionado no mapa.

A paleta Mapa (figura 3) permite a definição do tamanho e localização da área onde serão exibidos os mapas, o posicionamento e tipo dos elementos acessórios do mapa, como legenda e escala. Finalmente, a paleta Utilidades (figura 3) permite colocar funcionalidades adicionais no SIG Web como, por exemplo, botões de ajuda e atualização da página.

Independente do ambiente alvo desejado, a ferramenta SIGWeb Builder permite a criação de um SIG Web de forma visual e bastante interativa. O desenvolvedor do SIG Web pode agora concentrar seus esforços na estruturação e aparência do sistema e abstrair dos detalhes da implementação. Na verdade, o desenvolvedor SIG Web nem precisa conhecer estes detalhes ou possuir conhecimento prévio da tecnologia por trás do seu sistema.

Acreditamos que a ferramenta SIGWeb Builder seja uma importante contribuição na área de disseminação de informações geográficas na Web, não só por facilitar o desenvolvimento de SIG Webs, mas principalmente por permitir a popularização do trabalho de desenvolvimento desses sistemas.

4. Trabalhos Futuros

A ferramenta proposta neste trabalho está sendo implementada em Java, seguindo os conceitos de padrões de projeto. Esta ferramenta em seu estágio atual permite a criação de *layers* e mapas temáticos definidos nos Arquivos *XML* de Projeto do AlovMap e *MapFile* do MapServer. Para trabalhos futuros pretende-se:

- Geração de todos os arquivos do AlovMap e do MapServer, contemplando todas as funcionalidades destes ambientes;
- Inclusão do I3GEO com ambiente alvo [I3GEO 2008];
- Utilização da ferramenta em um estudo de caso no desenvolvimento de dois SIG Webs para a Empresa Baiana de Águas e Saneamento - EMBASA.

Gostaríamos de acrescentar que a ferramenta SIGWeb Builder adota a política de software livre e código aberto. Desta forma, poderá ser facilmente estendida e utilizada livremente.

Referências

AlovMap (2007). Disponível em: http://www.alov.org/. Último Acesso em: 10/11/2007.

Aragão, H. , Britto, D., et al (2007). "Desenvolvimento de um SIG para Saneamento usando Software Livre".24ºCongresso Brasileiro de Engenharia Sanitária e Ambiental-Minas Gerais.

Babu, M.N. (2003). "Implementing internet GIS with java based Client/Server Environment". Map Asia Conference. Kuala Lumpur, Malaysia.

Ceccarelli, M., Francesco, C. e Capua, M. (2006). "Dynamic WEBGIS and tracking of mobile devices". FOSS4G – Free and Open Source Software for Geoinformatics. Lausanne, Switzerlan.

I3GEO (2008). Disponível em: http://www.softwarepublico.gov.br/spb/ver-comunidade?community_id=1444332. Último Acesso em: 09/11/2007.

Kropla, B. (2005). "Begining Mapserver open source GIS Development", Apress.

MapServer (2007). Disponível em http://mapserver.gis.umn.edu/. Último acesso em: 18/12/2007.

Miranda, J. (2003). "Publicando mapas na web: Servlets, Applets ou CGI?". EMBRAPA, Campinas.

Peng, Z. e Tsou, M. Hsiang (2003). "Internet GIS: distributed geographic information services for the internet and wireless networks". New Jersey: John Wiley & Sons, Inc.

Tu, S. e Abdelguerfi, M. (2006). "Web Services for Geographic Information Systems". University of New Orleans. Published by the IEEE Computer Society.

Land change modeling and institutional factors: heterogeneous rules of territory use in the Brazilian Amazonia

Paulo Pimenta[1], Andréa Coelho[2], Sergio Costa[3], Evaldinólia Moreira[3], Ana Paula Aguiar [3], Gilberto Câmara[3], Roberto Araújo[4], Adagenor Ribeiro[1]

[1] Departamento de Informática, Universidade Federal do Pará (UFPA), Belém – PA

[2] Núcleo de Cenários, Instituto de Desenvolvimento Econômico-Social do Pará (IDESP), Belém – PA

[3] Centro de Ciências do Sistema Terrestre (CST), Instituto Espacial de Pesquisas Espaciais (INPE), São José dos Campos – SP

[4] Departamento de Antropologia, Museu Paraense Emílio Goeldi (MPEG), Belém – PA
pf.pimenta@gmail.com, andrea.coelho@idesp.pa.gov.br,
{scosta, eva, anapaula, gilberto }@dpi.inpe.br,
robertoa@museu-goeldi.br, adagenor.ribeiro@ufpa.pa

Abstract. *Land changes are determined by a complex web of biophysical and socio-economic factors that interact in time and space, in different historical and geographical contexts, creating different trajectories of change. It is people's response to economic opportunities mediated by institutional factors that drives changes. In this paper we discuss how to incorporate such institutional land tenure categories in land change models. Our hypothesis is that this is an essential step in the direction of constructing regional models to represent the heterogeneity of actors and processes. We implemented the conceptual proposal using the TerraME modeling Environment. Through a case study we analyze how the existence of different rules of territory use affects the landscape dynamics at the regional level.*

1. Introduction

Land change studies have a fundamental role in environmental research, since they establish a link between human activities and environmental systems. Land changes are determined by a complex web of biophysical and socio-economic factors that interact in time and space, in different historical and geographical contexts, creating different trajectories of change. Decisions that influence land change are made at different levels of organization (individual, household, community, nations, international/environmental trade agreements). It is people's response to economic opportunities *mediated by institutional factors* that drives changes [Lambin, 2001].

Modeling is one of the methods in the portfolio of techniques and approaches available to unravel the dynamics of the land use system [Verburg 2006]. There are different types of models in the literature, as reviewed in [Briassoulis 2000; Agarwal 2002; Verburg, Schot, Dijst et al. 2004; Verburg 2006]. Such models can be classified in many different ways, according to their goals and technical approaches. In this work, we analyze how to incorporate land tenure institutional factors in land change models.

The specific problem we face is to model land change in the Brazilian Amazonia. The Brazilian government, through a series of territorial planning policies in the last decades, created a mosaic of territory units and heterogeneous institutional contexts that strongly influence the land change dynamic in the region. Figure 1 illustrates the situation in Pará State.

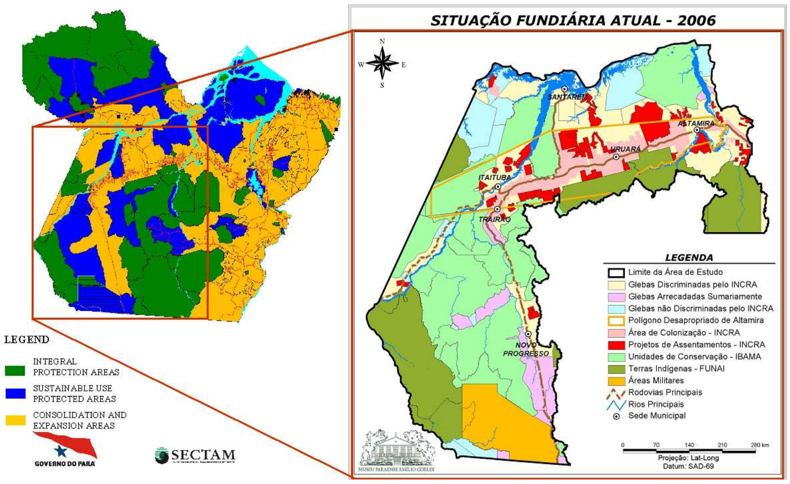


Figure 1. Example of mosaic of territory units in Pará State (sources: Pará State Government and Museu Emilio Goeldi).

These territory units may broadly be characterized into: (a) *Indigenous lands*; (b) Federal and State *Conservation Units* of several types, including the ones oriented towards biodiversity conservation, communitarian exploration resources, and wood exploration use concessions; (c) *Settlement Projects*, also of several categories, including those designated to familiar agriculture and, more recently, to agroextrativism exploration; (d) and, finally, *Expansion Areas*, which consist of all the areas available for expansion of cattle ranching, mechanized and familiar agriculture. Most activities in (d) occur in areas without regular titles of land property. These are, in general, the more conflicting areas in the Amazonia. Most of them still belong to the Federal and State governments, who have the power to discriminate the goal of a given unit. Territory units in categories (a), (b) and (c) have specific rules regarding their destination and land use. For example, in a recently created settlement modality called PDS (Sustainable Development Project), only a small number of families are favored, and their economic activities must be oriented towards forest products exploration. In this settlement modality, each family can clear-cut only 3 ha to subsistence agriculture. On the other hand, in (d), large, medium and small agriculture farmers must respect environmental rules according to the Federal Forest Law, which imposes that 80% of forest inside private properties must be preserved. In all categories, however, rules are often not followed by local actors.

Previous land-change modeling exercises in the region [Laurance, Cochrane, Bergen et al. 2001; Soares-Filho 2004; Aguiar 2006] did not explicitly considerer the fact that each of these categories has different actors and rules acting on them. In previous models, the existence of conservation units and indigenous lands slows down - or completely shields - forest conversion from happening. The heterogeneity of types of conversation units and settlement modalities were not considered. However, land changes *do* occur in the different categories, following specific rules regarding allowed actors, land uses and also speed of change. The scientific question we discuss in this paper is *how to incorporate such institutional land tenure categories and their specific rules in land change models.* Our hypothesis is that this is an essential step in the direction of constructing regional models to represent the heterogeneity of actors and processes, which could be used as tools to support public policies. Incorporating the main land tenure categories in regional models will help to make visible the “invisible actors and processes”, even when detailed cadastral information about individual properties are not available.

This paper is structured as follows. In Section 2, we present a brief revision about land-change models, focusing on the difference between *Top-down* and *Bottom-up* design approaches. Section 3 discusses conceptually how to incorporate the land tenure aspects in regional models. In Section 4, we describe a computational implementation of these concepts. Section 5 describes one case study in the Brazilian Amazonia, in Santarém, Pará State. Finally, Section 6 presents our conclusions.

2. Land change modeling overview

There is a large diversity of modeling approaches in the literature. In this review, we focus on a classification based on design approach: *Top-down* and *Bottom-up. Top-down* models originate from landscape ecology, and are based on remote sensing and census data. The *Top-down* approach describes a whole study area through of a statistical or mathematical formulation. On other hand, models conceived using a *Bottom-up* approach describe explicitly the actors of land change. They describe individual agents and their interaction with environment, as showed in Figure 2. Next section details and exemplifies both approaches.

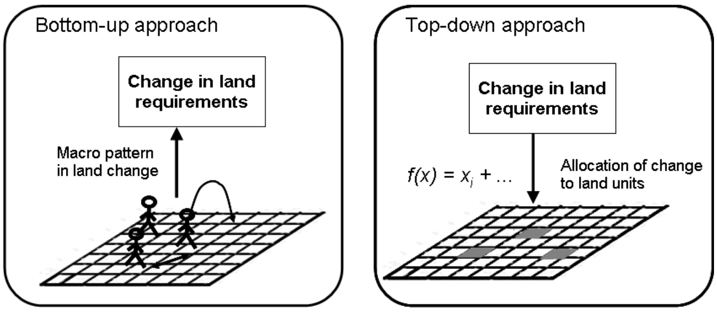


Figure 2. Top-down and bottom-up approaches (adapted from: [Verburg 2006]).

2.1 Top-down approach

Top-down models consider in general land-uses/covers as a set of discrete states. Land changes are transitions from one to another [Walker 2004]. Raster or cellular space, subdivided in pixels or cells represent each discrete state. This approach uses an empirical, mathematical, statistical or econometric equation to describe the transitions among states. There are different models in the literature: CLUE [Veldkamp and Fresco 1996; Verburg, De Koning, Kok et al. 1999], CLUE-s [Verburg, Soepboer, Veldkamp et al. 2002], Dinamica [Soares-Filho, Cerqueira and Pennachin 2002], RIKS [White and Engelen 1997; White and Engelen 2000], CA_Markov [Eastman 2003]. The structures of these models present some likenesses, as discussed in [Eastman, Solórzano and Fossen 2005] or [Verburg, Kok, Pontius et al. 2006]. Eastman [2005] argues that models consist of three major parts: a *change demand submodel, a transition potential* *submodel, and a change allocation submodel*. The *demand submodel* calculates the rate and magnitude of change, usually based on economic model, trend analysis, or scenario analysis to quantify the change (or demand). This demand is then allocated in a spatially explicit grid by the *change allocation*. This submodel uses a suitability (or change transition potential) maps representing the suitability/or potential for change of each cell for a given land use/or transition. This map is produced by *transition potential submodel*, given a set of input driving factors a method to relation these maps, as a multivariate statistical. The change allocation produces then a new land use map that can be use to next model iteration. In some cases, the rate of change may be modified by the results of the allocation module, in a bottom-up feedback mechanism [Verburg 2006], although this is not always the case. Most models adopt a purely top-down design. Figure 3 illustrates the structure of models using the top-down approach.

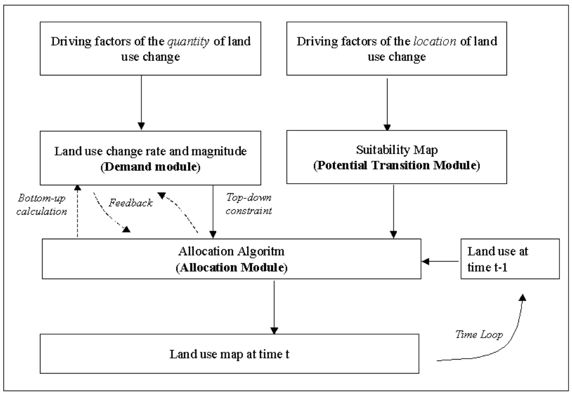


Figure 3 – The generalized model structure of spatially explict land-use change models. Source : [Verburg, Kok, Pontius et al. 2006].

2.2. Bottom-up approach

Bottom-up models are based in the concept of "emergence" that is an essential characteristic of social simulation, where the interactions among each entities at a lower level to result in macro pattern [Matthews, Gilbert, Roach et al. 2005]. General examples include bird-flocking model [Reynolds 1987] and segregation model [Schelling 1971]. In the land change context, this approach includes micro-economic and agent-based approaches. Microeconomic models consider that individual landowners make theirs decisions with the objective to maximize expected returns or utility derived from the land based in economic theory [Verburg, Schot, Dijst et al. 2004]. An agent-based model consists of autonomous entities (agents), an environment where the agents interact (normally represented by a cellular space), and rules that define the relations between agents and their environment [Parker, Berger, Manson et al. 2002]. Figure 4 shows the structure of agent-based models. Each agent can interact with other agents, and can modify the environment and influence other agents.

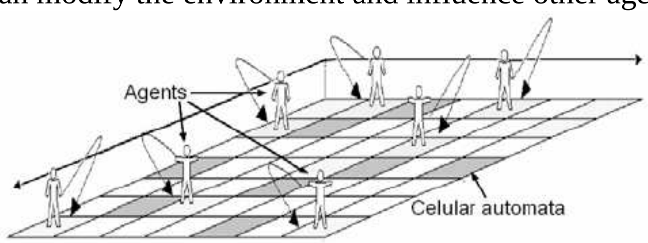


Figure 4. Agent based modeling: Adapted from [Parker, Berger, Manson et al. 2002].

2.3. Which approach is the best?

The use of either a top-down or bottom-up approach depends on the extent of analysis and the dominant processes of land use change in the studied area. The *Top-down* approach is adequate to processes in which the changes are largely steered by regional demand, as in the case of the expansion of cash crop agriculture in regions with abundant land resources [Verburg 2006]. Top-down models are easier and faster to construct than bottom-up models to larger areas. There are a number of available frameworks for top-down modeling (such as CLUE, CLUE-S, DINAMICA), which mostly requires parameterization for the specific case study, including the selection of driving factors, and definitions about how to compute demand and the transition potential. Remote sensing and census data can be use as the main inputs to analyze land use/cover patterns and derive driving factors relations. *Bottom-up* models require extensive fieldwork to design agents' rules and behavior, and are normally constructed for smaller areas, aiming at specific research questions. However, top-down models are derived on pattern analysis, and a series of simplifying assumptions are incorporated. They can't express the heterogeneity of actors and processes as well as Bottom-up/agent-based models ideally could. Bottom-up models have the potential to represent the complex biophysical and socioeconomic process, the interaction of actors at different levels of organization.

The selection of a given approach is very much dependent on the modeling goals. In the current *State of Art* of land change models, it is almost a tradeoff between heterogeneity and scale of analysis. One of the challenges of this scientific field is to combine both approaches, as purely bottom-up or top-down may be insufficient to represent biophysical and socioeconomic processes interactions at different levels of organization, from local to global. A discussion about hybrid approaches can be found in [Verburg 2006] and [Moreira, Costa, Aguiar et al. 2008].

In this paper, we focus mainly on Top-down land change models, as our main goal is the construction of regional scale models. Besides, models constructed using the Bottom-up approach would be naturally able to include such heterogeneity of actors and rules. In the case study presented in Section 5, we analyze how the existence of different rules of territory use affects the landscape dynamics at the regional level. In the next session, we present our conceptual proposal for the inclusion of such heterogeneous land use rules in top-down land-change models.

3. Different territory rules in top-down land change models: conceptual proposal

We analyze two aspects of the heterogeneity of rules in the Brazilian Amazonia in their incorporation in land change top-down models: (1) the temporal evolution of such rules, as different unit are created; (2) allowed uses and conversions in different units. As described in Section 2, Top-down models are composed of three sub-models: Demand, Allocation, Potential Transition. The incorporation of such heterogeneous rules can be achieved by the regionalization of one of these components, or a combination of them. Previous modeling works have done this in different contexts, trying to incorporate some level of heterogeneity in top-down models. In this section, we discuss some of the possibilities:

1. *Regionalize the three components*, creating different models for each unit. A similar approach was used in [Carneiro, Aguiar, Escada et al. 2004] with the goal of creating different models for large and small farmers. In our case, totally different models for each category could be implemented, including different Allocation and Transition Potential rules.
2. *Regionalize only the demand*: keep the same allocation and transition potential modules, but externally compute different speeds of change for each unit. Similar approach was used on Aguiar [2006] and Soares-Filho [2006] to force different rates of change in different regions of the Amazonia.
3. *Keep the three components basically the same*, and regionalize only the parameters for the potential transition module. A similar approach was used in some applications of the CLUE model in large regions such as China [Verburg and Veldkamp 2001].

The choice of a proper solution is a matter of the modeling goal. In the case study we discuss in this paper, our aim is to analyze the landscape dynamics for the whole area of study, considering the implications of the existence - and enforcement - of alternative rules. In this case, option 3 is more appropriated, as we are more interested in the distribution of change over the whole territory, than in actual amount of change. So, to illustrate such concepts, we implemented a top-down approach, with the same demand and allocation module, but using specific rules to compute the transition potential for different use in different categories of land tenure units. Such rules may vary along the time, as new units are created. We choose to adapt an existent top-down model, the CLUES-S [Verburg, Soepboer, Veldkamp et al. 2002]. The same adaptation could be implemented in similar models, like DINAMICA [Soares-Filho, Cerqueira and Pennachin 2002] and CLUE [Veldkamp and Fresco 1996].

4. Implementation

To test the concepts discussed in last section, we implemented a top-down model based on the CLUE-S framework [Verburg, Soepboer, Veldkamp et al. 2002] basic ideas and conceptual design. The new model was implemented using TerraME modeling environment [Carneiro 2006] CLUE-S is a typical top-down hierarchical model. The allocation module uses an externally defined *Demand*, computed using the usual trend analysis, scenario formulation, etc. The *Potential Transition module* computes the cell suitabilities for different uses using logistic regression analysis. Recent applications of the CLUE-S model modified the cell suitability computation to include expert knowledge. Castella and Verburg[2007] for example uses a rule-based version to CLUE-s model. The *Allocation Module* calculates, with discrete time steps, the most likely changes in land use given the suitabilities, and some possible restrictions, which may be informed as parameters to the allocation module:

- Allowed land use transitions in the study area (for instance, forest can be converted to agriculture, but not directly to secondary growth vegetation). These allowed transitions are represented as *Transition Matrices*.
- Spatial policies and restrictions, representing areas in which certain conversions cannot occur (for instance, forest conversion not allowed inside parks);

- Temporal restrictions indicating the minimum and maximum number of years before a conversion can occur. For instance, within a shifting cultivation system, the number of years a piece of land can be used due to soil nutrient depletion and weed infestation.

Another important parameter to the *Allocation module* are *land use elasticities*, which indicate the relative difficulty to modify a certain use (for example, a industrial area can be more difficult to remove and convert to another use than a pasture area).

The model we implemented is inspired and based on the concepts of the CLUE-s model as presented by [Verburg, Soepboer, Veldkamp et al. 2002]. The main difference relies in the possibility to extend the above-mentioned spatial restrictions regarding allowed conversions to allow multiple regions/transition matrices. In this paper, each region represents, a different type of territory unit, with their specific rules, as discussed in Section 1. In other applications, other types of regions could be emplyed. Another difference from the original CLUE-S resides in the fact that other parameters (land use elasticities and driving factors) can also be regionalized. The regionalization may evolve with time, as new territory units are created, through a *Regionalization Module*. Figure 5 illustrated these modifications. At each time step, the *Regionalization module* updates an attribute of each cell corresponding to the type of region it belongs. This attribute is used in the Allocation module to select the appropriate parameters (possible transitions represented as alternative Transition Matrices; Land Use Elasticities; and Driving factors) to compute the *Transition Potential* of each cell.

On the other hand, in comparison to the original CLUE-S model, our model does not include the temporal restrictions on the conversions (minimum and maximum number of years before a conversion can occur), nor recently added functionalities of the CLUE-S model, such as cellular automata /neighborhood functionalities.

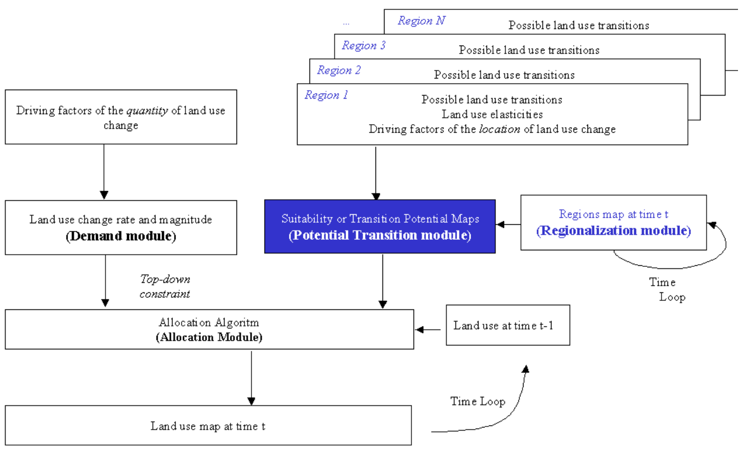


Figure 5. Modified top-down model to include regionalization.

5. Case Study

The goal of the modeling exercise we present here in to test the inclusion of the institutional land tenure factors in a real-world case study. We selected the Santarém region land use dynamics from 1999 to 2007. We compare the use of heterogeneous and homogeneous rules over the landscape, in order to analyze adherence of the official rules of use of the territory to the real patterns identified in the multi-temporal image classification, as described below.

5.1 Study area

The Santarém region, in Pará State, presented an intense land change dynamics due to the expansion of mechanized agriculture in the last decade, in substitution to familiar agriculture, secondary forest and forest areas. Figure 6 illustrates the study area. It comprehends part of the Santarém, Belterra and Placas municipalities.

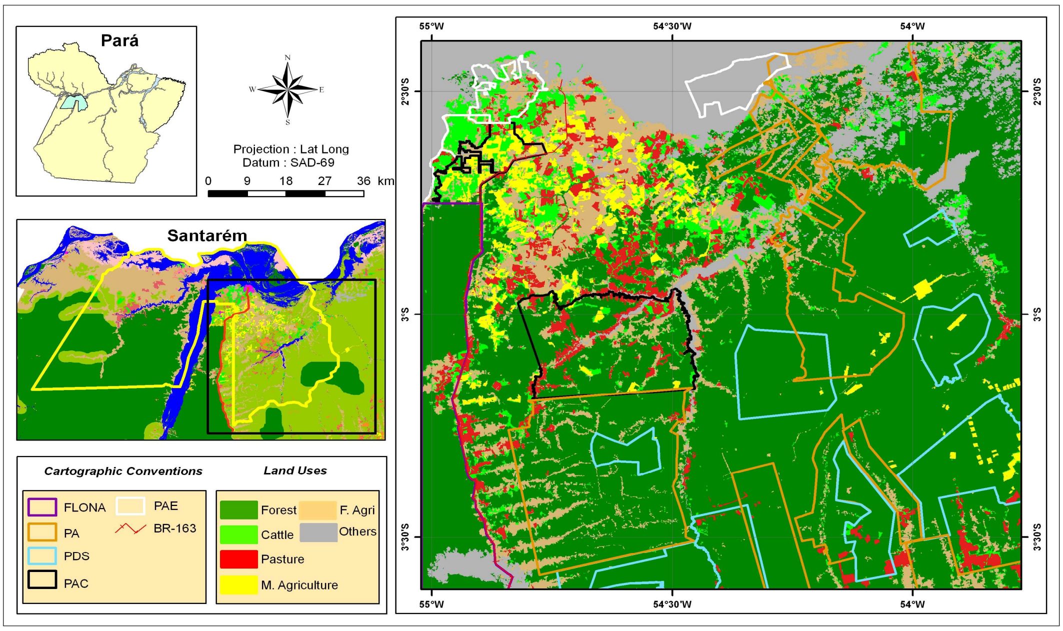


Figure 6 – Study Area (source: Andrea Coelho, IDESP).

As illustrated in Figure 6, this area encompasses a mosaic of different territory units, including a Conservation Unit, the Tapajós National Forest. A National Forest (FLONA) is a type of conservation unit that allows sustainable economic activities, according to a management plan for forest resources exploration. Existing population is allowed to stay when FLONAS are created, and practice subsistence agriculture. The study area also includes 27 Settlements of four different modalities:

- 10 PA (*Projeto de Assentamento*), designed to familiar agriculture, for purposes of colonization and homestead. It the traditional type of settlement created after the 80ths. Each family has less than 100 ha, which can be used to diverse agricultural activities, respecting the Federal Forest Law.
- 3 PAC (Projeto de Consolidação de Assentamentos): this modality was created to regularize the situation of already installed farms. Land property is collective.
- 4 PAE (Projeto de Assentamento Agro-extrativista), designed to traditional population, which should base their economy on extractive activities or agriculture. The property title is collective, and the use of natural resources has to be decided in a communitarian manner.
- 10 PDS (*Projeto de Desenvolvimento Sustentável*), designed to low environmental impact activities based on forest products. Each family is allowed to use only 3 ha to subsistence farming. Title is also collective.

FLONA Tapajós was created in 1974. The PA's in the study area were created after 1995. All the other units were recently created, after 2005. The official rules are not followed in many cases. As Figure 6 illustrates, in 2007 there is mechanized agriculture, which is associated to capitalized actors, inside PDS and PAC areas.

5.2. Spatio-temporal database

The first step of model construction is the creation of a spatio-temporal database for the study area. The database contains information about the land use dynamics and potential determining factors aggregated in regular cells of 500 x 500 m^2, representing three different dates: 1999, 2004 and 2007. The model considers five land use classes: *forest, secondary growth, cattle ranching, familiar agriculture, and mechanized agriculture*. They were obtained though classification of Landsat TM images. Cells representing water, cloud, and non-classified pixels were discarded from analysis. Driving factors of location of change include: distance to roads, to rivers, to urban centers, slope, and soils quality. The database also includes information about the different territorial units in the area. Each cell is classified according to the type, year of creation of the existing units in a given year.

5.3. Different transition rules according to institutional aspects

The relative importance of the driving factors was established using logistic regression analysis for each land use. These coefficients are used to compute cell suitabilities in the Transition Potential Module. These coefficients were adapted using field experience and expert knowledge, during a calibration phase. Elasticity factors were also determined during a model calibration phase, comparing model results to 2004 real patterns. In the results we show in this paper, we opted for maintaining the same driving factors, their suitability computation coefficients, and elasticity values for all types of units. The heterogeneity is translated into alternative Transition Matrices representing allowed land use conversion in each of the five types of units present in the study area: FLONA, PA, PAE, PDS, PAC and Expansion Areas. Table 1 illustrates such matrices.

PAC allowed Transition	Forest	Secondary Vegetation	Cattle Ranching	Mechanized Agricultur	Familiar Agricultur
Forest	1	0	0	0	1
Secondary Vegetation	0	1	0	0	1
Cattle	0	0	1	0	0
Mechanized	0	0	0	1	0
Familiar	0	1	0	0	1
Expansion Area Transition	Forest	Secondary Vegetation	Cattle Ranching	Mechanized Agricultur	Familiar Agricultur
Forest	1	0	1	1	1
Secondary Vegetation	0	1	1	1	1
Cattle	0	1	1	1	1
Mechanized	0	1	1	1	1
Familiar	0	1	1	1	1

Table 1. Examples of Transition Matrices for PAC and Expansion Areas

5.4. Exploration of alternative rules

In order to analyze the adherence of the official rules of use of the territory to the real land-use patterns dynamics in the period 1999-2004-2007, we compare the results of two alternative model simulations:

A. Using heterogeneous rules for each class of special area

B. Using a homogeneous rule, the same valid for the agriculture expansion area

Notice that the goal here is not to obtain the exact patterns (i.e., to achieve a high level of correctness in the classification), but to use the model as an additional tool to understand in which situations the rules are not being following, and to capture how the overall landscape dynamics is influenced by the mosaic of uses and special areas that compose the area. The demand for change (amount of change per land use) is obtained from the 2004 and 2007 maps, and interpolated for the intermediary years.

5.5. Simulation results in a PAC area

In this section, we illustrate some of the results obtained in the modeling process using the approach presented in this paper. We focus on a specific area in which rules are not being followed in a PAC area (Figure 7).

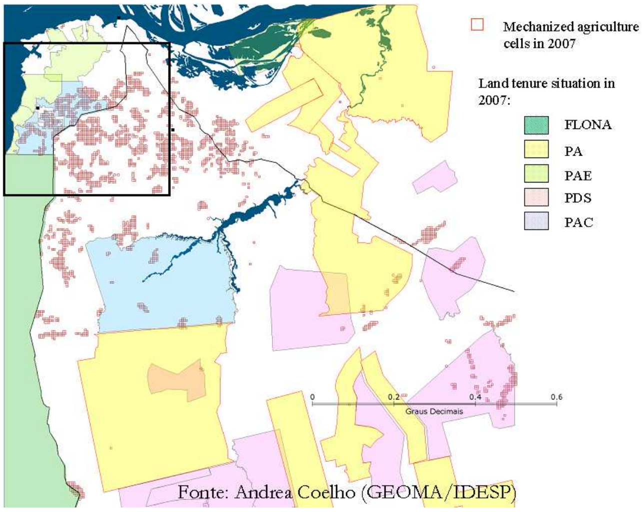


Figure 7– Selected area to analyze the territory rules enforcement in the land change model.

Figure 8.a and 8.b illustrate the real land change occurred in the area. Figures 8.c and 8.d compare two alternative results: considering or not the heterogeneous rules of territory use.

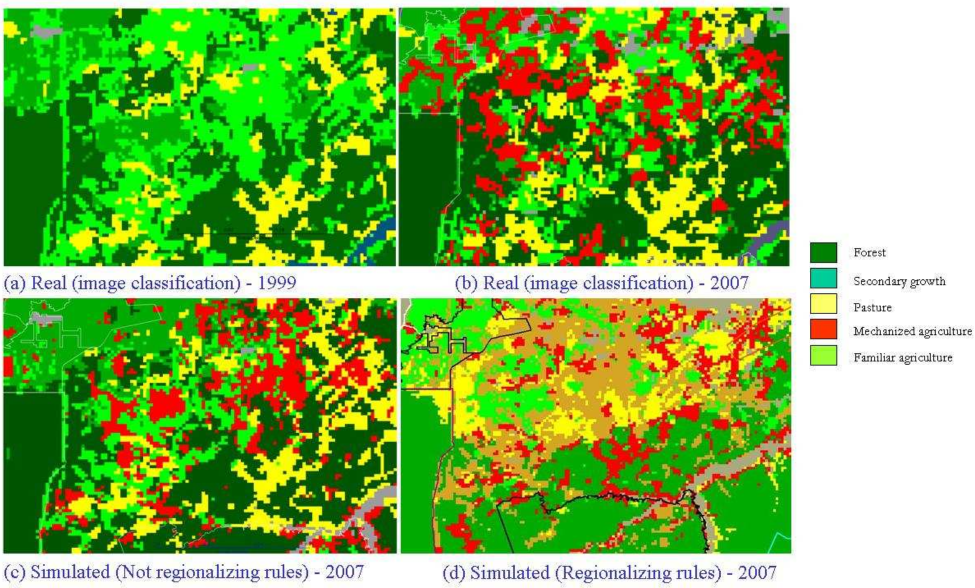


Figure 8– Results comparing PAC area: (a) Land use map in 1999; (b) Land use map in 2007; (c) Simulated results (2007) without regionalization; (d) Simulated results (2007) with regionalization.

As Figure 8.b illustrates, the PAC area is dominated by mechanized agriculture in 2007. Using the non-regionalized model, some of this process is captured in the simulated result (Figure 8.c). On the other hand, when the correct rule is applied to the PAC (Figure 8.d), a totally diverse result is obtained. The only allowed use, according to Table 1, is familiar agriculture.

6. Conclusion

In this paper, we discussed how to incorporate land tenure institutional land tenure categories and their specific rules in top-down land change models. We implemented a model to test our proposal in a agriculture expansion frontier in Central Amazonia. We conclude that the whole modeling process gets richer and insightful when such institutional aspects are considered. Land tenure information provides information about what we can (and can't) see in remote sensing derived information. Incorporating such information in regional models will help to make visible the "invisible actors and processes", even when detailed cadastral information about individual properties are not available.

Acknowledments

This work is part of the GEOMA Network Project (www.geoma.lncc.br). We thank the CLUE group in Wageningen University for the technological support during the development of our project. Paulo Fernando Pimenta has a CNPq scholarship, in the context of the PIME (Integrated Science and Technology Ministry and EMBRAPA Project). Gilberto Camara's work is partially funded by CNPq (grant PQ 550250/2005-0) and FAPESP (grant 04/11012-0). Evaldinolia Moreira's work is partially funded by FAPEMA (Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão).

References

Agarwal, C. G., Glen M.; Grove, J. Morgan; Evans, Tom P.; Schweik, Charles M. (2002). "A review and assessment of land-use change models: dynamics of space, time, and human choice."

Aguiar, A. P. D. (2006). Modeling land use change in the Brazilian Amazon: exploring intra-regional heterogeneity. São José dos Campos, INPE. **PhD:** 204.

Briassoulis, H. (2000). Analysis of Land Use Change: Theoretical and Modeling Approaches, Regional Research Institute, West Virginia University

Carneiro, T. (2006). Nested-CA: a foundation for multiscale modeling of land use and land change. PhD in Computer Science. São José dos Campos, National Institute of Space Research**:** 109.

Carneiro, T. G., A. P. Aguiar, M. I. Escada, et al. (2004). A Modelling Environment for non-isotropic and non-homogeneous spatial dynamic models development. International Workshop on Integrated assessment of the land system:The future of land use, Amsterdam, Wageningen University.

Castella, J.-C. and P. H. Verburg (2007). "Combination of process-oriented and pattern-oriented models of land-use change in a mountain area of Vietnam." Ecological modelling **202**: 410-420.

Eastman, J. (2003). IDRISI Kilimanjaro. Guide to GIS and Image Processing. Worcester, MA, Clark University.

Eastman, J. R., L. A. Solórzano and M. E. V. Fossen (2005). Transition Potential Modeling for Land-Cover Change. GIS, Spatial Analysis, and Modeling. D. Maguire, M. Batty and M. Goodchild. California, ESRI Press**:** 357-385.

Laurance, W., M. Cochrane, S. Bergen, et al. (2001). "The future of the Brazilian Amazon." Science **291**: 438-439.

Matthews, R., N. G. Gilbert, A. Roach, et al. (2005). "Agent-based land-use models: a review of applications." Landscape Ecology(Volume 22, Number 10 / December, 2007): 1447-1459.

Moreira, E. G., S. S. Costa, A. P. D. d. Aguiar, et al. (2008). Dynamic coupling of multiscale land change models: Interactions and feedbacks across regional and local deforestation models in the Brazilian Amazon. Landscape Ecology - Special Issue (submitted).

Parker, D., T. Berger, S. Manson, et al. (2002). Agent-Based Models of Land-Use /Land-Cover Change. Report and Review of an International Workshop. Irvine, California, USA, LUCC Project.

Reynolds, C. (1987). "Flocks, herds and schools: A distributed behavioral model." SIGGRAPH Comput. Graph. **21**(4): 25-34.

Schelling, T. (1971). "Dynamic models of segregation." Journal of Mathematical Sociology **1**: 143-186.

Soares-Filho, B., G. Cerqueira and C. Pennachin (2002). "DINAMICA – a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier." Ecological modelling **154**(3): 217 – 235.

Soares-Filho, B., D. Nepstad, L. Curran, et al. (2006). "Modeling conservation in the Amazon basin." Nature **4389**.

Soares-Filho, B. A., Ane; Nepstad, Daniel; Cerqueira, Gustavo; Vera Diaz, Maria del Carmen; Rivero, Sérgio; Solórzano, Luis; Voll, Eliane (2004). "Simulating the response of land-cover changes to road paving and governance along a major Amazon highway: the Santarém-Cuiabá corridor." Global Change Biology **10**(20): 745-764.

Veldkamp, A. and L. O. Fresco (1996). "CLUE: a conceptual model to study the Conversion of Land Use and its Effects." Ecological modelling **85**: 253-270.

Verburg, P. (2006). "Simulating feedbacks in land use and land cover change models." Landscape Ecology **21**(8): 1171-1183.

Verburg, P., G. De Koning, K. Kok, et al. (1999). "A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use." Ecological modelling **116**: 45-61.

Verburg, P. and A. Veldkamp (2001). "The role of spatially explicit models in land-use change research: a case study for cropping patterns in China." Agriculture, Ecosystems & Environment **85**(14): 177-190.

Verburg, P. H., K. Kok, R. G. Pontius, et al. (2006). Modelling land use and land cover change. Land-use and land-cover change. Local processes and global impacts, Springer, Berlim.

Verburg, P. H., P. P. Schot, M. J. Dijst, et al. (2004). "Land use change modelling: current practice and research priorities." GeoJournal **61**(4): 309-324.

Verburg, P. H., W. Soepboer, A. Veldkamp, et al. (2002). "Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model." Environmental Management **30**(3): 391-405.

Walker, R. (2004). "Theorizing Land-Cover and Land-Use Change: The Case of Tropical Deforestation." International Regional Science Review **27**(3): 247-270.

White, R. and G. Engelen (1997). "Cellular automata as the basis of integrated dynamic regional modelling." Environment and Planning B: Planning and Design **24**: 235-246.

White, R. and G. Engelen (2000). "High-resolution integrated modelling of the spatial dynamics of urban and regional systems." Computer, Environment and Urban Systems **24**: 383-400.

Spatial relations across scales in land change models

Evaldinolia G. Moreira[1, 2], Ana Paula Aguiar[1], Sérgio S. Costa[1], Gilberto Câmara[1]

[1]Earth System Science Center (CST) – Instituto Nacional de Pesquisas Espaciais (INPE) - Av. dos Astronautas, 1758 – 12227-001 – São José dos Campos – SP – Brasil

[2]Departamento de Informática – Centro Federal de Educação Tecnológica do Maranhão (CEFET-MA) – São Luis – MA - Brasil

{eva, anapaula, scosta, gilberto, }@dpi.inpe.br

***Abstract**: Land changes are the result of a complex web of interaction between human and biophysical factors, which act over a wide range of temporal and spatial scales. In this paper we conceptualize spatial relations among geographic objects at different scales. We analyze two types of spatial relations: hierarchical, which handles the interaction of nested objects at different scales (countries, states and municipalities, for example); and a network-based relation, which handles action-at-a-distance types of interaction (market chains, for instance) of objects at different scales (such as farms in Central Amazonia to soybean market consumers at the global scale). We implemented such relations in the Terralib environment, in which they can be constructed using selected strategies, and then used in dynamic models. We exemplify the use of such concepts in a real-world case study in the Brazilian Amazonia. We conclude that combining hierarchical and network-based spatial relations provide a comprehensive conceptual framework to include top-down and bottom-up interactions and feedbacks in multi-scale land-change models.*

1. Introduction

Modelling land change involves the use of representations of interactions within the land use system to explore its dynamics and possible developments [Verburg, Eickhout and Meij 2008]. Models can also be used to project the impact of policy changes on the current land use trajectory [Pijanowskia, Brownb, Shellitoc et al. 2002]. Land changes are the result of a complex web of interaction between human and biophysical factors, which act over a wide range of temporal and spatial scales. At different scales of analysis different factors have a dominant influence on the land use system: at the micro scale, land use patterns may be determined by household structure and local biophysical constraints. At the regional level the distances to markets and regional climate variations may determine land use pattern. Regional dynamics impact and are impacted by local dynamics through *top-down* and *bottom-up* interactions [Verburg, Schot, Dijst et al. 2004]. Understanding processes of change from the local to the global scale and their impacts on the coupled human-environmental system is a main scientific challenge [Moran, Ojima, Buchmann et al. 2005].

Land change processes are also intimately linked to processes of globalization. Globalization is the growing and accelerated interconnectedness of the world in an economic, political, social and cultural sense. It increasingly separated places of consumption from places of production, such that land systems cannot be adequately understood without knowing their linkages to decisions and structures made elsewhere. In this sense, understanding the role of networks is essential to understanding land-use structure [Verburg, Schot, Dijst et al. 2004]. Such networks can be physical, such as infrastructure networks, and logical ones, such as market chains, linking a certain location to distant consumption or influential sites. According to Becker [2005]: "*it is impossible today, more than ever, to understand what happens in one place without considering the interests and conflicting actions at different geographical scales*".

The goal of this paper is to discuss the incorporation of such hierarchical and network spatial relations in multi-scale land change models. We consider no single model or scale can handle the complexity of interactions that influence land change. Multi-scale land change models have been developed to address these issues. Some multi-scale modelling approaches combine different spatial models at different scales, mostly simulating *top-down* influences [Verburg, Eickhout and Meij 2008]. *Bottom-up* interactions and scaling issues have started to be addressed by multi-agent systems [Parker, Berger, Manson et al. 2002], in which interactions among individuals can simulate the emergent properties of the systems. Most land use change modelling embody the notion of space as a set of absolute locations in a Cartesian coordinate system, thus failing to incorporate spatial relations dependent on topological connections and network fluxes. Current land change models often deal with spatial interactions over large regions using (transport) network analysis to compute driving factors representing travel times and distant to ports, markets, etc. In spite of the progress in multi-scale modelling and spatial interaction analysis, there is still a need for approaches and techniques to deal adequately with scaling interaction issues [Verburg, Kok, Pontius Jr et al. 2006]. Understanding the interactions between and across scales, and the effects of globalization on local land-change processes, will remain the research frontier of land use/land cover for the next decade.

Our work contributes in this direction. We conceptualize spatial relations among geographic objects at different scales. We analyze two types of spatial relations: *hierarchical*, which handles the interaction of *nested* objects at different scales (countries, states and municipalities, for example); and *action at a distance* that handles the interaction through a *network* (market chain, for instance) of objects at different scales (such as grid cells in Central Amazonia to wood market consumers at the global scale). We argue these spatial relations provide a comprehensive conceptual framework to include *top-down* and *bottom-up* interactions and feedbacks in multi-scale models. This paper is organized as follows. Section 2 discusses the conceptual definition of these multi-scale spatial relations, and presents the implementation of these concepts using the Terralib GIS library [Câmara, Souza, Pedrosa et al. 2000]. Section 4 exemplifies, using a real world case study in the Brazilian Amazonia, how the explicit definition of such *hierarchical* and *action at a distance spatial relations* allow the representation of *top-down* and *bottom-up* linkages in multi-scale models.

2. Spatial relations across scales in land change models

In this paper, we use the definition of scale given by Gibson et al [2000]: "*scale is the spatial, temporal, quantitative, or analytical dimension used to measure and study any phenomenon*". All scales have extent and resolution. In the case of spatial scales, extension refers to the dimension of the study area, and resolution to the measurement precision. At each spatial scale geographic objects may be differently represented. Examples of representation of such objects include: (a) area regions whose boundaries are closed polygons; (b) cellular automata organized as sets of cells, whose boundaries are the edges of each cell; (c) point locations in two-dimensional space. For simplicity, we refer to the representation of spatial objects as *Entities*.

Our goal is to conceptualize the spatial relations between pairs of *Entities* at different scales to allow a broad representation of *top-down* and *bottom-up* interactions in land change models. We discuss two types of spatial relations: hierarchical and relative space relations.

Several existing land change models are organized in *top-down* manner, in which a demand for change is spatially allocated according to cell suitability. This includes the above mentioned CLUE [Veldkamp and Fresco 1996] and CLUE-S [Verburg, Soepboer, Veldkamp et al. 2002], Dinamica [Soares, Cerqueira and Pennachin 2002] GEOMOD [Pontius, Cornell and Hall 2001], and Environmental Modeler [Engelen, White and Nijs 2003]. Such models use *Hierarchical spatial relations*, in which nested scales are combined, as exemplified in Figure 1. The Environmental Modeler uses three different scales. Economic models at the national and regional scales compute land requirements for different land uses, based on economic and demographic factors. These land requirements are then allocated in a regular grid using a cellular automata model at the local scale. The CLUE model framework consists of two components: a demand module, that projects the amount overall of change; and an allocation module, the spatial component that acts in two scales (a coarse and a fine resolution grid) to localize such changes, based on cell suitability.

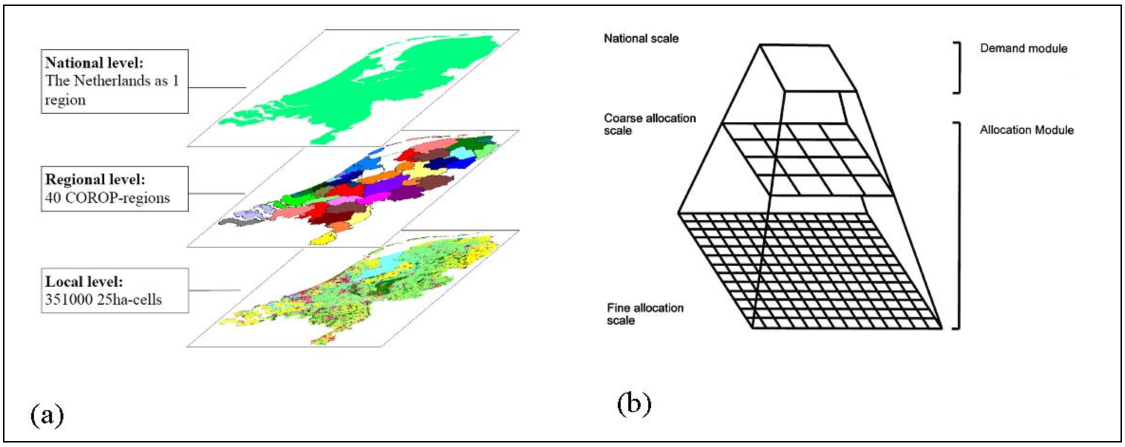


Figure 1. Examples of hierarchical structures used in land-change models: (a) Environmental Modeler [Engelen, White and Nijs 2003]; (b) CLUE model [Veldkamp and Fresco 1996].

Entities in these cases are regular cells with different resolution at different scales, or polygons representing administrative units at different levels of organization. The spatial relations represent parenthood relations (father-son and son-father). Father-son hierarchical relations are necessary to inform lower-scale model of the context provided by higher-level models, and are the most common type of relation found in current land-change models. Son-father relations allow local scale models to inform regional models (*bottom-up* interactions). Although to some extent *bottom-up* interactions have been included in hierarchical land-change models (for example, Verburg et al. [Verburg, De Koning, Kok et al. 1999]), the full integration of *top-down* and *bottom-up* scale interactions is still a research topic [Verburg, Kok, Pontius Jr et al. 2006].

Such hierarchical spatial relations embody the notion of space as a set of absolute locations in a Cartesian coordinate system. However, flows of resources, information, organizational interaction and people are essential components of space, and should be treated in land change models. Efficient representation of such flows in connection with representation of absolute space is essential to achieve a realistic perspective of spatial relations, and inform land-change models [Harvey 1989]. These flows, which are normally represented as networks [Aguiar, Câmara and Cartaxo. 2003; Verburg, Schot, Dijst et al. 2004; Verburg, Kok, Pontius Jr et al. 2006], link processes that act on different scales. The global and continental market connections in Amazonia are an example of this, as Figure 2 illustrates. Different flows connect areas in the region to distance places of consumption, influencing the land use system in heterogeneous ways. Wood products from Brazil are mostly exported to Europe, as Figure 2.a illustrates. However, internal market also plays an important role in the wood market. Becker (2001) estimates about 80% is sold to the Southeast of Brazil. Global markets play a determining role for other commodities too. Santarém, in Pará State, has international connections related to the international soybeans markets, due to the presence of Cargill in the area. São Felix do Xingu, also in Pará, has different national and international connections related to the meat market, due to the presence of global companies like Bertin. The IRSSA (South-American Regional Infra-structure Integration Initiative, [IIRSA]) integration axes (Figure 2.b) will change the commercial connectivity of places like Roraima and Amapa, due to the Guiana-Venezuela-Suriname planned axe (Guiana Shield Hub). Large container transport companies, such as CMA-CGM, have already announced they will use the Madeira River corridor to export wood, cotton, and meat. The Madeira corridor is also part of Brazilian Infrastructure Plans for the Amazonia, linking Porto Velho, Rondonia State, to Manaus, in Amazonas State. Incorporating such heterogeneous connections in land change models is essential to improve our understanding about their impacts on the land use system, and to envision the future scenarios for the region.

Combining such hierarchical and network-based relations is necessary to provide the necessary conceptual support to multi-scale land change models. Sections 2.1 and 2.2 present a conceptualization of these two types of relations. Our implementation of such concepts is briefly described in Section 2.3.

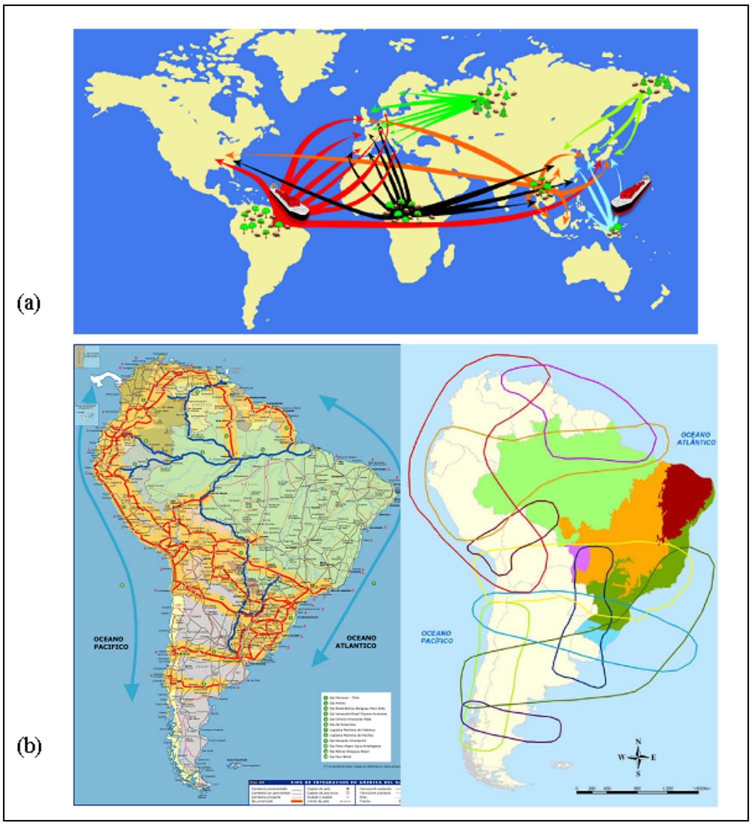


Figure 2. Examples of links to global and continental markets: (a) International flow of wood from Amazonia (source: Greenpeace, www.greenpeace.org); (b) IIRSA infra-structure integration axes in South America, facilitating the commercial flow of different areas to Europe, North-America and/or Asia markets.

2.1. Hierarchical relations

We propose to represent a hierarchical relation as a directed graph *G* composed of set of nodes N_1 and N_2, representing *Entities* at $Scale_1$ and $Scale_2$; and a set of arcs *A* linking nodes N_1 to N_2.

The arcs *A* can have attributes or not, depending on the strategy used to construct them. When *Entities* at both scales have an area representation (polygons or regular cells), we propose three alternative strategies, illustrated in Figure 3. They are based on topological relations as described below.

- *Simple*: when spatial resolutions are perfectly matched, simple "within" or "coveredby" or "equals" spatial operator can define the parenthood relation between scales.

- *ChooseOne*: for area representations, when hierarchical spatial resolutions do not match, this strategy chooses the upper scale unit cells with larger percentage of intersection as the father and the "intersection" spatial operator can define the relation.

- *KeepInBoth*: also only for area representations, when hierarchical spatial resolutions do not match, this strategy keeps all intersected upper units cells as fathers and the “intersection” spatial operator can define the relation. The percentage of each intersection is stored as an attribute of the Arc *A*.

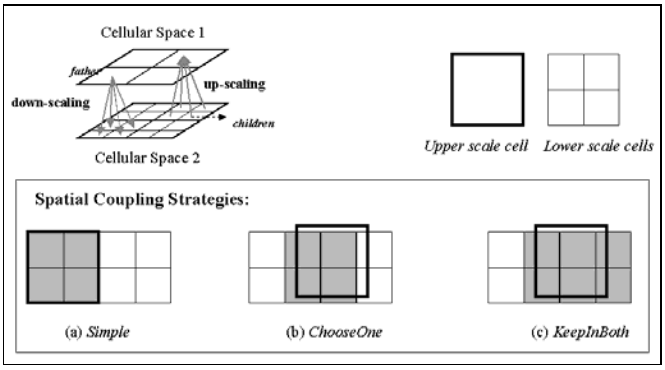


Figure 3. Schematic representation of strategies for spatial coupling in the case of regular cells: (a) Simple; (b) ChooseOne; (c) KeepInBoth.

Hierarchical networks can represent spatial relations of point entities at different scales, such as urban centers (State capital at the macro scale; major cities at the meso scale; villages at the local scale). To construct graph *G* in this case, manual or attribute based strategies could be envisioned (for example, administrative unit name to establish son-father relations). The attributes of the Arcs *A* could also be derived from geographical objects (such as percentage of population).

2.2. Network-based relations

We also represent *network-based relations* as a directed graph G composed of set of nodes E_1 and E_2, representing *Entities* at $Scale_1$ and $Scale_2$, and a set of arcs *A* linking nodes E_1 to/from E_2. The representation is the same as for hierarchical relations. The difference resides in the strategies to construct *G*. A network *T* is required to represent physical (roads, rivers, energy) and logical (airline routes, market chains, migration fluxes) linkages between elements E_1 and E_2. These linkages will be established using network analysis operators.

According to characteristics of the network, specific construction strategies will decide: (a) if two nodes in E_1 and E_2 are connected; (b) the strength of this connection. The construction strategies presented here are based on the concepts introduced by Aguiar et al [2003] regarding the construction of a *Generalized Proximity Matrix* (GPM). The GPM represents *absolute and relative space neighborhood relations* among objects of the same type, at the same scale. A GPM is used to support spatial analysis and cellular automata dynamic models. We modify the GPM construction strategies to consider objects of different types, at different scales to support the development of multiscale land-change models. Two strategies are then proposed:

- *Multi-scale Closed-networks* linkages: to connect entities at different scales using networks in which the entrances and exits are restricted to its nodes. They encompass *logical* (such as banking networks and productive chains) and some types of *physical networks* (railroads, telecommunication networks).

• *Multi-scale Open-networks* linkages: to connect entities at different scales using networks in which in which any location is entrance or exit point. These are always *physical networks*. Examples are transportation networks such as roads and rivers. For open networks, it is necessary to make use of the actual line coordinates that correspond to each arc in order to be able to compute the closest entrance/exit points from any arbitrary position.

The strategies can be summarized as follows:

• For each object in O_1, compute the nearest entry point E_1 in network *T*.

• For each object in O_2, compute the nearest entry point E_2 in network *T*.

• The existence of a linkage from E_1 to/from E_2 is computed using network analysis.

Figure 4 illustrates the process of constructing graph *G* to represent relative space relations. A set of parameters bounds connectivity limits according to network and case study characteristics. For instance, one can define objects at $Scale_1$ are not linked to the network if they are more than a 100 km away from the closest entry point. Limits can also be imposed for minimum path in the network. For instance: only objects at $Scale_1$ not more than 10 hours from the markets (represented at $Scale_2$) through the infrastructure network are considered connected. Minimum path computation depends on network attributes. Different case studies can use, for example, distance or travel time (infrastructure networks), flow of people (migration networks), dollars (banking networks), added value (production chains).

Note that when *Entities* at both scales have an area representation (polygons or regular cells), the connection is performed using the area centroid.

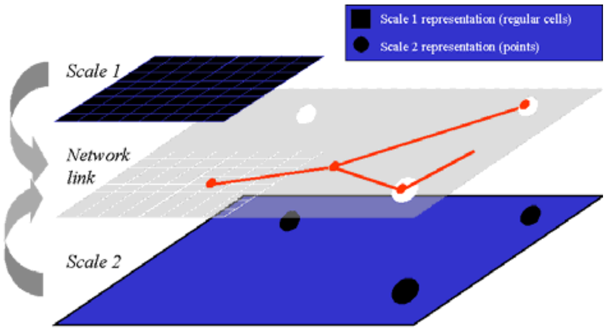


Figure 4. Schematic representation of a network-based spatial relation between cell objects in Scale 1 and point objects in Scale 2.

2.3. Implementation

We implemented the conceptual definitions presented in Sections 2.1 and 2.2 using the Terralib GIS library. For both types of relations, construction strategies were added to the library as an extension of the Generalized Proximity Matrix functionality [Aguiar, Câmara and Cartaxo. 2003]. The new strategies deal with relations among objects in two different layers of information, representing the geographic objects at different scales. The relations can be constructed for polygon, points and cell representations of objects. For hierarchical relations only the absolute space Cartesian coordinates are considered to define the father-son and son-father relations. For network-based relations, a third layer is necessary representing a logical or physical network used to define the connectivity between the objects. The case study described below uses the Terralib implementation to construct the graphs *G* representing the relations. They are stored in a database, and can be exported as text files. Once constructed, tools for dynamic modelling can be applied using the relations. We use the TerraME modelling environment [Carneiro 2006] to develop our land-change models using the stored relations.

3. Example in the Brazilian Amazonia

We exemplify the use of concepts presented in the previous sections in a multiscale land change model for the Brazilian Amazonia developed by Aguiar [2006] and Moreira et al [2008], using the TerraME modelling environment [Carneiro 2006].

The model encompasses three scales: (a) at the national level, the main markets for Amazonia products (Northeast and São Paulo) and the roads infrastructure network; (b) at the regional level, a regular grid of 25 x 25 km^2 resolution cells for the whole Brazilian Amazonia, covering an area of approximately 4 million km^2; and (c) at the local level, a nested regular grid of 1 x 1 km^2 resolution cells for a hot-spot of deforestation in Central Amazonia, the Iriri region, in São Felix do Xingu, Pará State. This local grid covers an area of approximately 50,000 km^2. Figure 5 illustrates the three scales and their geographic objects representation.

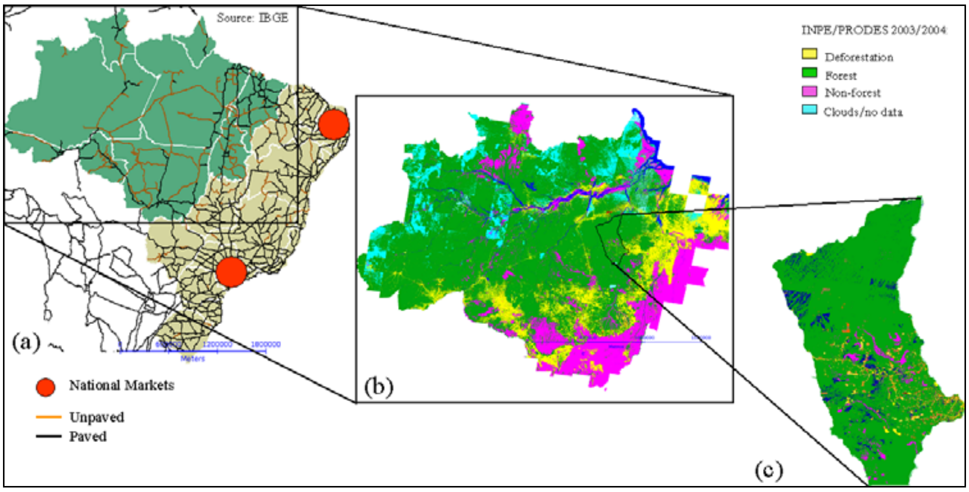


Figure 5. Study area: (a) Brazil: roads network and main markets (São Paulo and Northeast); (b) Brazilian Amazonia: deforested areas map; (c) Iriri/Terra do Meio in São Felix do Xingu, Pará State: deforested areas map.

The goal of the multi-scale model is explore *complementary information about the occupation process* in the region [Moreira, Costa, Aguiar et al. 2008]. The model includes the following interacting model components and spatial relations:

- The regional scale model projects the percentage of deforestation for each 25 x 25 km^2 cells. We used a statistical allocation procedure based on regression models adapted from the CLUE model [Veldkamp and Fresco 1996] by Aguiar [2006]. It represents the process of agricultural frontier expansion over the whole Brazilian Amazonia. The macro model seeks to answer questions such as:

given a certain pressure for expansion of agricultural land, which areas in the Amazonia would be occupied first? One of the goals was to explore the hypothesis that connection to national markets through roads infrastructure is a key factor to explain the distribution of deforestation in the region. *This requires the establishment of a network-based relation to link the cells in Amazonia to places outside the region.* This relation is used in defining the suitability of the 25 x 25 km^2 cells for change according to their level of connectivity.

- The nested local model seeks to answer questions such as: given that a certain amount of pressure is projected for the Iriri by the regional model, how would local patterns of occupation evolve? The *top-down* interaction consists of the regional model signalling an expected demand for change at the Iriri. This demand is calculated as a function that sums-up the deforested area (father-cells) at the regional scale and sends it to the local scale. *This requires a father-son relation to select the 25 x 25 km^2 cells corresponding to the Iriri 1 x 1 km^2 cells.* The model uses this relation to add the large-scale projected change at 25 x 25 km^2 cells and send the resulting a demand for change to the local model.

- The Iriri model is an agent-based deforestation model [Moreira, Costa, Aguiar et al. 2008]. Two sets of agents were identified: small and large farmers. Small settlers favour proximity to roads and urban centres. Large farmer prefer large pieces of inexpensive land, not necessarily close to roads. Therefore, each type of actor is associated to set of determining factors and decision rules. Local policy decisions, expressed at local scale, may prevent the full extent of projected change from occurring. A *bottom-up feedback mechanism* sends this information back to the larger scale and thus modifies the macro scale model corresponding cells. *This requires a son-father relationship to link 1 x 1 km^2 cells to the upper-scale 25 x 25 km^2 cells.* The model used this relation to correct the projected change at the 25 x 25 km^2 cells.

To support the implementation of such scale interactions in this land-change model, we defined and computed the following hierarchical and network-based relations.

3.1. Hierarchical relation between the nested grids

We used a hierarchical relation to provide the spatial support to dynamically link the two nested grids at 25 x 25 km^2 and 1 x 1 km^2 resolutions. The strategy we use to construct the relation is the *KeepInBoth,* as the cellular spaces were not coincident.

Each coarse scale cell is linked to approximately 625 finer scale cells (father-son relation). Most finer scale cells are linked to only one coarser scale cells (son-father relation), but depending on their relative position (on the borders of the coarse scale cells) they can be linked to two, three or even four parent cells (see Figure 3.c). The father-son and son-father hierarchical relations allow the incorporation of *top-down* and *bottom-up* interactions between the regional and local models, as discussed in Section 2.1.

3.2. Network-based relation: connection to markets

We used an *Multi-scale Open-network* strategy to connect the regional scale 25 x 25 km^2 cells to the main places of consumption at the national scale (São Paulo and Northeast). Graph *G* representing the relation between these objects is computed using the following parameters:

- Maximum distance from cells to the road network: unbound (all cells are included).
- Maximum distance from entrance points *E* through the network: unbound.
- Weight computation: inversely proportional to the minimum path distance from the cell to each national market, using the roads network. We distinguished paved from non-paved roads (non-paved roads are supposed to double the distances).

Graph *G* includes the *2:n* relationship from the two markets to every cell, and a *n:2* relationship from every cell to the two markets. Both directions could be used in land-change models. For example, the *2:n* (from market to cells) could be used to establish a remote influence between São Paulo and their most connected cells. We could include a rule in the model to bound change in Amazonia cells as a result of a behavioural or policy change in São Paulo. This change in the market conditions can be an incentive (demand increase) or a restriction (need of certification).

In this paper, the land-change model uses the *n:2* relationship (from cell to market). We derive a new cell attribute based on graph G to represent the level of connectivity of each cell to any of the markets. If road conditions change, the variable is recomputed. Each cell receives as attribute *conn_markets* the minimum *weight* value stored in G according to the roads network at that time. Figure 6 illustrate the connection to markets variable in 1997 and the projected 2010 level of connectivity, supposing some roads are to be paved.

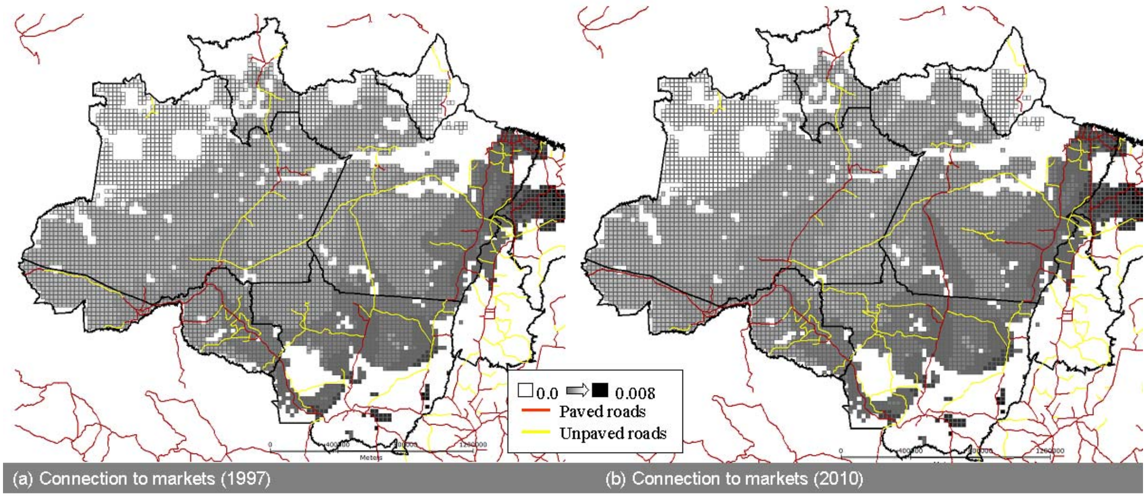


Figure 6. Connection to markets variable constructed using a network-based multiscale spatial relation: (a) in 1997; (b) in 2010 (paving some roads).

This network-based relation was used to construct one of the main variables in the model. Using the connection to national markets, the model was capable of reproducing the different stages of occupation of the new frontiers of the XXI century, using 1997 as the base year, including São Felix do Xingu [Aguiar 2006] comparing to 2003 deforestation maps (INPE, 2008). The model captures the process in which cattle ranchers decided to migrate to the São Felix area due to its biophysical, accessibility and market conditions. The connection to markets variable represents a process that acts in a higher hierarchical level, and could not be captured in a single scale study.

4. Conclusions

This paper discussed and conceptualized the use of multi-scale spatial relations in land change models. Two types of relations were presented: hierarchical and network-based. Multi-scale land-change models are often based on hierarchical relations, using nested objects at different scales. We argue that combining hierarchical relations with network-based relations provide a comprehensive conceptual framework to include *top-down* and *bottom-up* interactions and feedbacks in multi-scale land-change models. Network-based relations can represent remote influences in the land use system. This has a growing importance in a globalized economy, in which places of consumption and production are increasingly separated. Land systems cannot be adequately understood without knowing the linkages of different areas to decisions and structures made elsewhere. We exemplified the use of such relations in a multi-scale land change model for the Brazilian Amazonia.

Acknowledgements

We thank the Terralib team for the support in the development of the GPM extensions, and the CLUE group in Wageningen University. This work is part of the GEOMA Network Project (www.geoma.lncc.br). Gilberto Camara's work is partially funded by CNPq (grant PQ 550250/2005-0) and FAPESP (grant 04/11012-0). Evaldinolia Moreira's work is partially funded by FAPEMA (Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão).

References

Aguiar, A., G. Câmara and R. Cartaxo. (2003). Modeling Spatial Relations by Generalized Proximity Matrices. V Brazilian Symposium in Geoinformatics - GeoInfo 2003, Campos do Jordão, SP, Brazil.

Aguiar, A. P. D. (2006). Modeling Land Use Change in the Brazilian Amazon: Exploring Intra-Regional Heterogeneity. PhD Thesis, Remote Sensing Program. Sao Jose dos Campos, INPE. PhD.

Câmara, G., R. Souza, B. Pedrosa, et al. (2000). TerraLib: Technology in Support of GIS Innovation. II Brazilian Symposium on Geoinformatics, GeoInfo2000, São Paulo.

Carneiro, T. (2006). Nested-CA: a foundation for multiscale modeling of land use and land change. PhD Thesis in Computer Science (avaliable at

www.dpi.inpe.br/gilberto/teses/nested_ca.pdf). Computer Science Department. Sao Jose dos Campos, INPE. Doctorate Thesis in Computer Science.

Engelen, G., R. White and T. d. Nijs (2003). "Environment Explorer: Spatial Support System for the Integrated Assessment of Socio-Economic and Environment Policies in the Netherlands." Integrated Assessment 4(2): 97-105.

Gibson, C. C., E. Ostrom and T. K. Ahn (2000). "The concept of scale and the human dimensions of global change: a survey." Ecological Economics 32(2): 217-239.

Harvey, D. (1989). The Condition of Postmodernity. London, Basil Blackwell.

IIRSA. (2005). "Initiative for Integration of Regional Infrastructure in South America." 2008, from http://www.iirsa.org.

Moran, E., D. Ojima, N. Buchmann, et al. (2005). Global Land Project: Science Plan and Implementation Strategy. IGBP Report No. 53/IHDP Report No. 19. Stockholm, IGBP Secretariat.

Moreira, E. G., S. S. Costa, A. P. D. d. Aguiar, et al. (2008). Dynamic coupling of multiscale land change models: Interactions and feedbacks across regional and local deforestation models in the Brazilian Amazon. Landscape Ecology. Special Issue (*Submitted*).

Parker, D., T. Berger, S. Manson, et al. (2002). Agent-Based Models of Land-Use /Land-Cover Change. Report and Review of an International Workshop. Irvine, California, USA, LUCC Project.

Pijanowskia, B. C., D. G. Brownb, B. A. Shellitoc, et al. (2002). "Using neural networks and GIS to forecast land use changes: a Land Transformation Model." Computers, Environment and Urban Systems 26: 553–575.

Pontius, R. G., J. Cornell and C. Hall (2001). "Modeling the spatial pattern of land-use change with GEOMOD2: application and validation for Costa Rica." Agriculture, Ecosystems & Environment 85(1-3): 191-203.

Soares, B. S., G. C. Cerqueira and C. L. Pennachin (2002). "DINAMICA - a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier." Ecological Modelling 154(3): 217-235.

Veldkamp, A. and L. Fresco (1996). "CLUE: A Conceptual Model to Study the Conversion of Land Use and its Effects." Ecological Modeling 85: 253-270.

Verburg, P., G. De Koning, K. Kok, et al. (1999). "A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use." Ecological Modelling 116: 45-61.

Verburg, P. H., B. Eickhout and H. v. Meij (2008). "A multi-scale, multi-model approach for analyzing the future dynamics of European land use." Annals of Regional Science 42(1): 57 - 77.

Verburg, P. H., K. Kok, R. G. Pontius Jr, et al. (2006). Modeling Land-Use and Land-Cover Change. Land-use and land-cover change: local processes and global impacts. E. Lambin and H. Geist. Berlin, Springer.

Verburg, P. H., P. P. Schot, M. J. Dijst, et al. (2004). "Land use change modelling: current practice and research priorities." GeoJournal 61(4): 309-324.

Verburg, P. H., W. Soepboer, A. Veldkamp, et al. (2002). "Modeling the Spatial Dynamics of Regional Land Use: The CLUE-S Model." Environmental Management 30(3): 391-405.

Irregular Cellular Spaces: Supporting Realistic Spatial Dynamic Modeling over Geographical Databases

Tiago G. S. Carneiro[1], Gilberto Câmara[2], Raian V. Maretto[1]

[1] Computer Science Department – Federal University of Ouro Preto, MG, Brazil

[2] Image Processing Division – National Institute for Space Research, SP, Brazil

{tiago, raian}@iceb.ufop.br, gilberto@dpi.inpe.br

Abstract. *This paper presents a new computational model for representing the Geographic Space concept, called Irregular Cellular Space, which main goal is supporting the development of multiscale spatially explicit dynamic models integrated to geographical databases. This model has been implemented in a modeling software platform named TerraME which has been used for the development of some interesting environmental dynamic models and form simulation of dynamic spatial patterns of change.*

Resumo. *Este artigo apresenta um novo modelo computação para representar o conceito de Espaço Geográfico, chamado Espaço Celulares Irregulares, que tem como principal objetivo servir como suporte para o desenvolvimento de modelos dinâmicos espacialmente explícitos integrados a bancos de dados geográficos em múltiplas escalas. Este modelo foi implementado em um ambiente de modelagem chamado TerraME que tem sido utilizado para o desenvolvimento de modelos ambientais dinâmicos e para a simulação de padrões espaciais de mudança.*

1 Introduction

The modern Geographic Information Systems (GIS) lack on representing dynamical aspects from geographic space. The most provides only a static computational representation for geo-objects, geo-fields and fluxes. This fact had led to several proposals of integration between dynamical modeling and GIS platforms [Box 2002] [Villa and Costanza 2000] [North et al. 2006]. In general, the space have been represented as a **regular cellular space** (RCS), i. e., a regular two-dimensional grid of multi-valued cells grouped into neighborhoods, where the dynamic model rules operate and possibly change cells attribute values. Due to the regular structure of this spatial model, the resulting modeling platforms inherit all disadvantages from raster representations for the space concept: border effects and cell attributes aggregation problems dependent on the chosen grid resolution [Costanza and Maxwell 1994] [Kok and Veldkamp 2001], and the lacking of abstractions for representing moving objects, as flights and trucks following their routes, or geographical networks, as roads or communication lines. Figure 1 illustrates some of these disadvantages, border effects and aggregation of cell attribute values.

There are several reasons for highlighting cellular spaces as a promising space representation for computer based dynamic modeling and simulation. The existence of a simple and formal cellular space based model of computation, which concept and structure may be efficiently specialized for dynamic space representation on discrete environments, the Cellular Automata (CA) [von Neumann 1966] model. In a Euclidian two-dimensional grid, one may use basic analytic geometry knowledge to describe change circular or elliptical paths. It is easy to develop algorithms for representing process trajectories: to go to East just increment the X coordinate, to go to the South decrement the Y coordinate. However, besides its simplicity, the CA model has enough complexity to simulate spatial diffusive processes and emergent phenomena [Batty 1999] [Wolfram 1984].

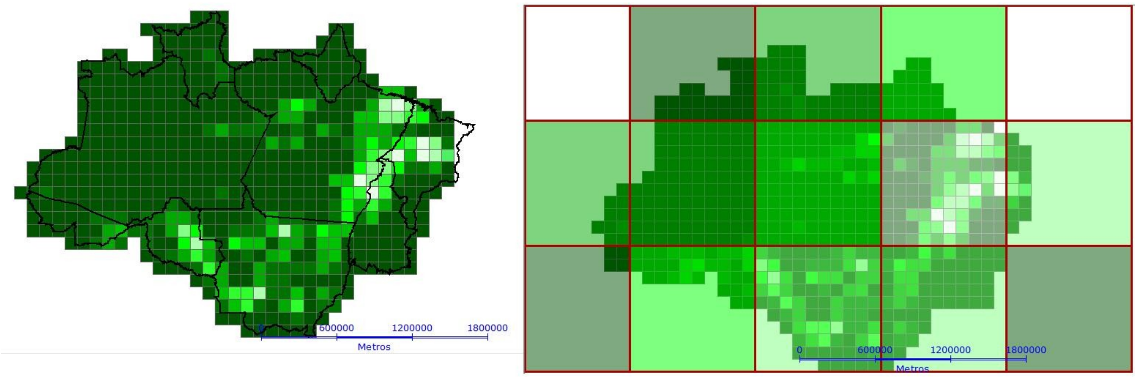


Figure 1. Problems due to the choice of a raster structure for space representation: aggregation of cell attribute values and border effects. Maps color: light green means “deforested” and dark green means “forest”.

This work presents a formal model for the geographical space concept, called **Irregular Cellular Space** (ICS), which extends the spatial structure from the RCS to support the development of GIS integrated spatial dynamic models which uses many space representations for supporting multiple scale modeling. For model evaluation, the ICS conceptual model has been implemented in environmental modeling software platform called TerraME. Its properties have been fully stressed in some interesting use cases

[Aguiar et al. 2005] [Almeida et al. 2008]. This paper exercises these properties in a simple and pedagogic deforestation model for the Brazilian Amazon region. Figure 2 shows the three ICS used in the deforestation model described at the end of this paper.

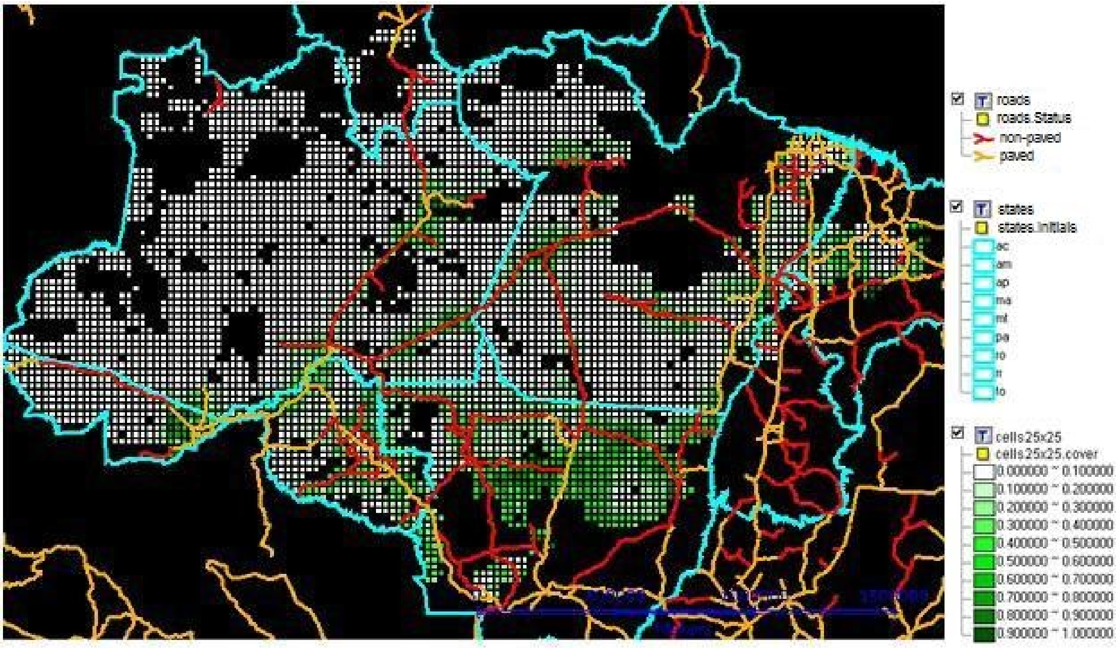


Figure 2. Three Irregular Cellular Spaces: (1) 25x25km2 sparse squared cells which main attribute is landCover *(white = "100% forest" and green = "0% forest")*; (2) each polygon representing one Brazilian State is a cell which main attribute are *name = {"MG" | "SP" | " RJ" | ...| "AM"}* the demanded area to be deforested; and (3) each roads is a cell which main attributes are status *(red = "paved" and orange = "non-paved")* and *brazilianSate = {"MG" | "SP" | " RJ" | ...| "AM"}.*

2 Basic Definitions

Following Castells (1999), this work views geographical space as a combination of "spaces of fixed locations and spaces of fluxes", where the concept of 'spaces of fixed locations' represents the geographical space as arrangements of objects on the **absolute space** (location), and the concept of 'spaces of fluxes' indicates spatial arrangements based on **relative space** (situation). Couclelis (1997) has proposed the notion of **proximal space** which aims to formalize this point of view.

In the regular cellular space model, a proximal space model, cells are indexed by their two-dimensional coordinates (x, y) meaning their relative position in the lattice of cells, and the cells neighborhood relationships are stationary following the Moore or Von

Newman pattern [Couclelis 1997]. In the ICS model there is no rigid structure for the space representation. The cellular space is any irregular arrange of cells which geometrical representation may vary from a regular grid of same size squared cells to a irregular set of points, lines, polygons, nodes and arcs, pixels, or even voxels. The n-dimensional space elements are indexed by a family of modeler defined functions, named **spatial iterators**, which map cell index values onto cells references. Topological relationships are expressed in terms of **Generalized Proximity Matrixes** (GPMs) allowing the representation of non-homogenous spaces where the spatial proximity relations are non-stationary and non-isotropic. A GPM is a weighted matrix where the spatial relations are computed taking into account both absolute space relations such as Euclidean distance or adjacency and relative space relations such as topological connection on a network [Aguiar and Câmara 2003].

3 The Irregular Cellular Space Model

(definition 1) The ICS is a set of cells defined by the structure *(S, A, G, I, T)*, where:

- $S \subseteq R^n$ is an n-dimensional Euclidian space which serves as support to the cellular space. The set $\boldsymbol{S}$ is partitioned into subsets, named cells, $\boldsymbol{S} = \{S_1, S_2, \ldots, S_m \mid S_i \cap S_j = \emptyset, \forall i \neq j, \cup S_i = S\}$.
- $A = \{(A_1, \leq), (A_2, \leq), \ldots, (A_n, \leq)\}$ is the set of partially ordered domains of cell attributes, and where a_i is a possible value of the attribute $(A_i, \leq)$, i.e., $a_i \in (A_i, \leq)$.
- $G = \{G_1, G_2, \ldots, G_n\}$ is a set of GPMs – Generalized Proximity Matrix (Aguiar, Câmara et al. 2003) used to model different non-stationary and non-isotropic neighborhood relationships, allowing their use of conventional relationships, such as topological adjacency and Euclidian distance, but also relative space proximity relations, based, for instance, on network connection relations.
- $I = \{(I_1, \leq), (I_2, \leq), \ldots, (I_n, \leq)\}$ is a set of domains of indexes where each $(I_i, \leq)$ is a partially ordered set of values used to index cellular space cells.
- $T = \{T_1, T_2, \ldots, T_n\}$ is a set of spatial iterators defined as functions of form $T_j : (I_i, \leq) \rightarrow S$ which assigns a cell from the geometrical support S to each index from $(I_i, \leq)$. Spatial iterators are useful to reproduce the spatial patterns of change since they permit easy definition of trajectories that can be used by the model entities to traverse the space applying their rules. For instance, the distance to urban center cell attribute can be sorted in an ascendant order to form

an index set *(I_i, ≤)* that, when traversed, allows an urban growth model to expand the urban area from the city frontier.

3.1 Spatial Iterators: modeling spatial trajectories of change

(definition 2) A spatial iterator $T_i \in T$ is an function defined as $T_i:(I_i, \leq) \rightarrow S$ that maps modeler built partially ordered sets of index $(I_i, \leq) \in I$ into cells $s_i \in S$.

The following functions should be defined by the modeler in order to construct the set of indexes $(I_i, \leq)$ and later uses it to build a spatial iterator.

- **(definition 2.1)** $filter: Sx(A_i,\leq) \rightarrow Boolean$ is a function used to filter the ICS, selecting the cells that will form the spatial iterator domain. It receives a cell $s_i \in S$ and the cell attributes $a_i \in (A_i, \leq)$ as parameters and returns "true" if the cell s_i will be inserted in $(I_i, \leq)$ and "false" if not.
- **(definition 2.2)** $\leq:(Sx(A_i,\leq))x(Sx(A_i,\leq)) \rightarrow Boolean$ is the function used to partially order the subset $(I_i, \leq)$ of cells. It receives two cell values as parameters and returns "true" if the first one is greater than the second, and otherwise it returns "false".
- **(definition 2.3)** $SpatialIterator: SxAxRxO \rightarrow T$ is a constructor function that creates a spatial iterator value $T_i \in T$ from instances of functions of the families R and O, where R are the filter functions as in definition 2.1 and O are the ≤ function as in definition 2.2. The SpatialIterator function is defined as: $SpatialIterator(filter, \leq) = \{(a_i, s_i) \mid filter(s_i, a_i) = true\ \forall a_i \in (A_i, \leq)$ and $\forall s_i \in S; a_i \leq a_j\ \forall i \leq j; s_i = spatialIterator(filter, \leq)\ \forall s_i \in S$ and $a_j \in (A_i, \leq)$ where $i = j\}$.

Figure 3 shows a source code piece, written in the TerraME modeling language, where a spatial iterator is created in order to simulate the deforestation process in Land Use and Cover Change (LUCC) models. The first parameter is the ICS for which the spatial iterator "it" is being created. The filter function, second parameter, selects only cells form "cs" whose land cover is "forest". The function ≤, third parameter, orders the cells according to their distance to the nearest road, making cells closer to roads more suitable to change. To construct a spatial iterator that traverses a two-dimensional ICS according to its Euclidian coordinates, from North to the South and from West to the East, one may define the function ≤ as: $\leq(c1, c2) = \{$ "true" if $c1.x < c2.x$; or "false" if $c1.x > c2.x$; or $(c1.y < c2.y)$ otherwise$\}$.

```
it = SpatialIterator{
        cs,
        function( cell ) return cell.cover == "forest"; end,
        function( c1, c2 ) return c1.distRoad > c2.distRoad; end
}
```

Figure 3. A TerraME representation of a spatial iterator used for modeling a deforestation process which spreads along the roads

3.2 Dynamic Operations on ICS

Different operations have been defined for traversing the ICS space or the cells neighborhoods applying rules which may cause changes: **ForEachCell**, **ForEachNeighbourhood**, and **ForEachNeighbor**.

- **(definition 3)** ForEachCell:TxF→A denotes the function that uses the spatial iterator $T_i \in T$ to traverse an ICS applying a modeler defined function $f_m \in F$, where F is the family of functions from the form f_m:SxNxA→A that calculates the new values for the attributes $a_j^t \in A_j$ from the cell $s_j \in S$ received as parameter. These functions also receives two others parameters: $n \in N$ a natural number corresponding to the relative cell position in the partially ordered set $(I_i, \leq) \in I$ used to define the spatial iterator T_i, and $a_j^{t-1} \in A$ the old values of the attributes a_j^t .
- **(definition 4)** ForEachNeighbourhood:SxGxF→A is a function which traverses the set of neighborhoods, G, from the cell received as parameter and applies a modeler defined function $f_v \in F$ to each cell neighborhood $g_i \in G$, where F is the family of functions from the form f_v:G→Bool. The function f_v receives a neighborhood g_i as parameter and returns a Boolean value: true if the ForEachNeighbourhood function should keep traversing the cell neighborhoods, or false if it should stop.
- **(definition 5)** ForEachNeighbor:SxGxF→A is a function which receives three parameters: a cell $s_i \in S$, a reference to one of neighborhood $g_i \in G$ defined for this cell, and a function $f_n \in F$, where F is the family of functions from the form f_m:(SxA)x(SxA)xR→Bool. The ForEachNeighbor function traverses the

neighborhood g_j and for each defined neighborhood relationship it applies the function f_m with the parameters $f_m(s_j, s_j, w_{ij})$, where $s_j \in S$ is the s_i neighbor cell and w_{ij} is a real number representing the relationship weight.

3.3 Application on Land Use and Cover Change Modeling

LUCC models distinguish between the projections for the quantity of change and for the location where these changes will take place [Veldkamp and Lambin 2001]. First, a sub-model which has rules that govern the amount of change (the "how much?" question) runs. It is called "demand model". Then, another called "allocation model" determines where the projected change will take place (the "where?" question). This structure is shown in Figure 4. At the next step, the LUCC models are back to the first stage until the simulation finishes.

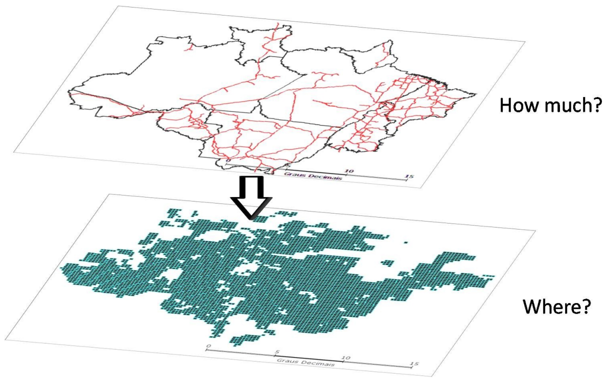


Figure 4. The general structure of LUCC model, with two sub-models: the "demand model" and the "allocation model".

A pedagogic LUCC model which simulates the deforestation process in Brazilian Amazon region has been developed. It stresses ICS properties for supporting multiple scales modeling through the use of multiple computational representations for the space concept. Three ICS models have been used to represent the deforestation process at two different scales, Figure 1. The demand model uses the cellular space composed by Brazilian States (polygons) and the other formed by Brazilian roads (lines). Each State has two attributes {deforestDemand, forestArea}. Let S_{BR} be a Brazilian State. At each simulation step (year), the amount of area demanded for deforestation in the State SBR is calculated as "SBR.deforestDemand = realRate * SBR.foresArea", where "realRate" is the annual deforestation rate, which is calculated as "realRate = absRate * paved", where “paved” is the percentage of paved roads in the State SBR and “absRate” is the absolute deforestation rate, which is directly proportional to the density of roads in the State SBR. The equation used to calculate this rate is shown at Figure 5, where “SBR.kmRoads” is the sum of the perimeter of all the roads of the state SBR, “SBR.area” is the area of the state SBR, “totalKmRoads” is the sum of the perimeter of all the roads of all states, “totalArea” is the sum of the areas of all states and “deforestRate” is a model parameter provided by the modeler which represents the average deforestation rate for the whole Legal Amazon area. As the roads change their status dynamically (Figure 7), that is, at each four years ten percent of the perimeter of the roads from each State changes its “status” attribute value from “non-paved” to “paved”, then the States deforestation rates are also dynamic. Initially, all cells were 100% forest, in other words, the value of each “landCover” cell attribute was equal to the cell area ($25x25km^2$). The total forest area of a State can be calculated as “S_{BR}.forestArea = S_{BR}.forestArea - S_{BR}.deforestDemand”.

$$absRate = \frac{\frac{S_{BR}.kmRoads}{S_{BR}.area}}{\frac{totalKmRoads}{totalArea}} * deforestRate$$

Figure 5. Equation used to calculate the absolute deforestation rate.

The location model uses only the sparse cellular space of small squared cells to determine where the changes will take place. It is based on a common approach: to compute a change potential surface. At this approach, each small squared cell will have a numeric value indicating how prone it is to change (deforestation). Then the model traverses the cellular “surface” in an ascending order of potential, applying the changes [White et al. 1998]. Some LUCC models use multi-linear regression for change potential computation, such as the CLUE model [Veldkamp and Fresco 1996]. Other approaches include a stochastic combination of diffusive and spontaneous change, such as the DINAMICA model [Soares et al. 2002]. In other to reproduce the “fishbone” spatial pattern of land occupation, where the deforestation spreads along the roads, the change potential of each cell has been computed through the spatial iterator defined in Figure 3, which resulting change potential surface has been shown in Figure 6(a). To simulate the land occupation process which is based only on the spatial expansion of the old human settlements, the change potential should be computed as in Figure 6(b). The surface on Figure 6(b) has been defined by the function ≤ (c1, c2) = {c1.distUrban < c2.distUrban}, where “distUrban” is the cell distance to the nearest urban center. Figure 8 shows the model results for the first, fifth and tenth years of simulation.

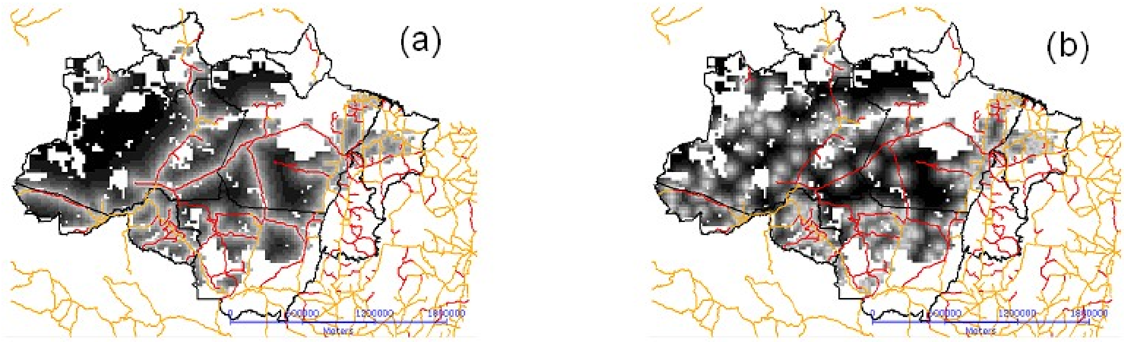


Figure 6. Change potential surfaces defined through spatial iterators based on (a) distance to road or (b) distance to urban centers. The gray scale surface reflects the potential for change of each cell: dark gray means low potential for change and light gray means high change potential.

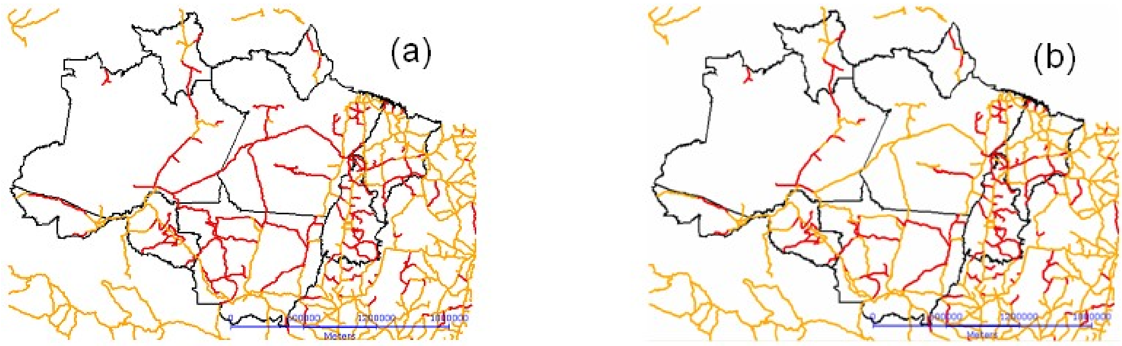


Figure 7. In central Amazon area: non-paved roads (red) in the past (b) become paved (orange) in the future (b).

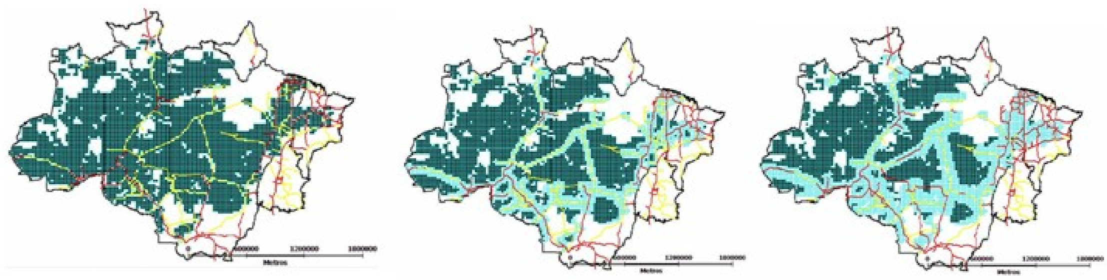


Figure 8. Deforestation model results: dark green are the 100% forest cells and light blue are the 100% deforested cells.

4 Related Works

To understand the needs related to the space representation in GIS integrated modeling platforms, we examine the proposed extensions of the CA model on the LUCC modeling literature. Several theoretical papers have proposed CA extensions for a better representation of geographical phenomena (Couclelis 1997; Takeyama and Couclelis 1997; Batty 1999; O'Sullivan 2001). In the specific case of LUCC modeling, recent works extend the original CA model and make it more suitable for representing the complexity of human-environment interaction (White, Engelen et al. 1998; Soares, Cerqueira et al. 2002; Almeida 2003). However, most extensions are not useful for multiple scale spatial dynamic modeling.

As an alternative for single-scale modeling of environmental changes, some authors have proposed the layered CA model (Straatman, Hagen et al. 2001), where every cell in one layer has one parent cell in the upper layer and an arbitrary number of child cells in the lower layer. This arrangement allows the combination of models that operate in different spatial resolutions. However, the layered CA model requires a decision about the spatial stratification, where each cell is dependent on a parent cell and controls a number of child cells. The layered CA falls short of providing adequate support for multiscale modeling, since it handles only layers of fixed spatial resolutions. This approach constrains the generality of the system, since the different processes are constrained to fit the hierarchical spatial structure.

5 Results and Future Works

The contributions from this work can be divided in three parts:

(a) The formal model for representing the Geographic Space concept: the ICS allows realistic multiple scale dynamic modeling through the simultaneous use of different computational representations for the space concept;

(b) The TerraME spatial dynamic modeling platform: the ICS model has been implemented in a software platform, named TerraME, which supports GIS integrated environmental model development;

(c) ICS applications: the ICS model has been used in some important modeling studies in order to simulated the human-environment interaction in the Brazilian Amazon region [Aguiar et al. 2005] and in some Brazilian National Parks [Almeida et al. 2008].

Many problems in the multiple scales spatial dynamic modeling has been not addressed on this work. However, the ICS model may be a first step towards a computational model for representing dynamic spaces on GIS environments. Among the future works, many other operations can be defined for the ICS model, for instance, operation to create neighborhood relationships or operations to couple cellular spaces form different resolutions, i. e, to couple two different scales.

6 References

Aguiar, A. P. D., Kok, K., Câmara, G., Escada, I. (2005) "Exploration of patterns of land-use change in the Brazilian Amazon using the CLUE framework." Proceedings of the Open Meeting of the Human Dimensions of Global Environmental Change Research Community, 6. Bonn, Germany.

Aguiar, A. P., Câmara, G., Monteiro, A. M., Souza, R. C. (2003) Modeling Spatial Relations by Generalized Proximity Matrices. Proceedings of Brazilian Symposium in Geoinformatics, 2003.

Almeida, R. M., Macau, E. E. N., França, H., Ramos, F.M. (2008) "Modelo de propagação de fogo em incêndios florestais e a teoria de percolação", XXXI National Conference on Applied and Computational Mathematics, Brazil.

Batty, M. (1999) "Modeling urban dynamics through GIS-based cellular automata". Computers, Environment and Urban Systems, v. 23, p.205-233.

Box, P. W. (2002) "Spatial units as agents: Making the landscape an equal player in agent-based simulations. In: Gimblett, H. R. (ed). Integration of agent-based modelling and geographic information systems. London UK: Oxford University Press, 2002.

Castells, M. (1999) "A Sociedade em Rede.", São Paulo: Paz e Terra.

Costanza, R. and T. Maxwell (1994)"Resolution and Predictability: an Approach to the Scaling Problem." Landscape Ecology v. 9, no. 1, pp 47-57.

Couclelis, H. (1997) "From cellular automata to urban models: New principles for model development and implementation". Environment and Planning B-Planning & Design, v. 24, n. 2, p. 165-174.

Kok, K.; Veldkamp, T. (2001) “A. Evaluating impact of spatial scales on land use pattern analysis in Central America.” Agriculture Ecosystems & Environment, v. 85, n.1-3, p. 205-221.

North, M.J., Collier, N.T., Vos, J.R. (2006) "Experiences Creating Three Implementations of the Repast Agent Modeling Toolkit," ACM Transactions on Modeling and Computer Simulation, Vol. 16, Issue 1, pp. 1-25, ACM, New York, New York, USA.

O'Sullivan, D. (2001) Graph-cellular automata: a generalised discrete urban and regional model. Environment and Planning B-Planning & Design, v. 28, n. 5, p. 687-705, 2001.

Soares, B. S., Assunção, R. M. (2001). “Modeling the spatial transition probabilities of landscape dynamics in an Amazonian colonization frontier”. Bioscience, v. 51, n. 12, p. 1059-1067.`

Straatman, B., Hagen, A. (2001). “The Use of Cellular Automata for Spatial Modelling and Decision Support in Coastal Zones and Estuaria” M. M. T. R. I. f. K. a. Systems. Maastricht, The Netherlands: Maastricht University.

Takeyama, M.; Couclelis, H. (1997) “Map dynamics: Integrating cellular automata and GIS through Geo-Algebra” International Journal of Geographical Information Science, v. 11, n. 1, p.73-91.

Veldkamp, A.; Fresco, L. O. (1996) CLUE: a conceptual model to study the Conversion of Land Use and its Effects. Ecological Modelling, v. 85, p. 253-270.

Veldkamp, A.; Lambin, (2001) E. F. “Predicting land-use change”. Agriculture Ecosystems & Environment, v. 85, n. 1-3, p. 1-6.

Villa, F., Costanza, R. (2000) .Design of multi-paradigm integrating modeling tools for ecological research Environmental Modeling and Software, Elsivier, Volume 15, Issue 2, pp.169-177

Von Neumann, J.. (1966) “Theory of self-reproducing automata”. Illinois: A.W. Burks.

White, R.; Engelen, G. (1998). „Vulnerability assessment of low-lying coastal areas and small islands to climate change and sea level rise – Phase 2: Case study St. Lucia. Kingston, Jamaica: United Nations Environment Program” - Caribbean Regional Coordinating Unit.

Wolfram, S. (1994) “Cellular automata as models of complexity.” Nature, v. 311, p. 419-424.

Desenvolvimento de uma Plataforma Gráfica para a Descrição de Modelos de Sistemas Ambientais

Tiago F. M. Lima[1,2], Tiago G. S. Carneiro[2], Sérgio D. Faria[3]

[1]Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais Instituto de Geociências – Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte, MG – Brasil

[2]TerraLAB - Departamento de Computação – Universidade Federal de Ouro Preto (UFOP)
Ouro Preto, MG – Brasil

[3]Instituto de Geociências – Universidade Federal de Minas Gerais (UFMG)
Belo Horizonte, MG – Brasil

`tiagofml@yahoo.com.br, tiago@iceb.ufop.br, fariamaracai@yahoo.com.br`

Abstract. *This paper presents the actual stage of development of TerraME Graphical Interface for Modeling and Simulation – TerraME GIMS, a graphical interface to building dynamic spatial-temporal models to the platform TerraME. The development of models through graphical components improve the productivity of the users and become more intuitive the process of building models, contributing to expand the user community of the environment TerraME.*

Resumo. *Este trabalho apresenta o estágio atual de desenvolvimento do TerraME Graphical Interface for Modeling and Simulation – TerraME GIMS, uma plataforma gráfica para a construção de modelos dinâmicos espaço-temporais para o ambiente de modelagem e simulação TerraME. A construção de modelos através de componentes gráficos além de aumentar a produtividade dos usuários torna o processo de construção de modelos mais intuitivo, contribuindo para ampliar a comunidade de usuários do ambiente TerraME.*

1. Introdução

Muitos problemas de natureza complexa exigem um tratamento interdisciplinar, envolvendo pessoas das mais diversas áreas do conhecimento. A criação de um modelo para representar o processo de mudança de uso e cobertura do solo na região Amazônica poderia envolver biólogos, sociólogos, economistas, geógrafos, geólogos, tecnólogos etc. A representação de um modelo desta natureza deve ser de comum entendimento para todas as pessoas envolvidas, sendo específica o suficiente para facilitar a representação do conhecimento que cada um tem a respeito do sistema e geral o suficiente para não restringir essa representação a um domínio de aplicação específico. A escolha da linguagem a ser utilizada na representação de um modelo é, portanto, fator crucial para o sucesso de um projeto de modelagem que envolva a colaboração de profissionais de diferentes formações.

O *TerraME* (Carneiro 2003) é um ambiente de modelagem e simulação de processos ambientais com representação explícita no espaço que permite aos usuários criar as estruturas, dados e regras, que irão definir o comportamento de um modelo, utilizando a linguagem de programação *TerraME*. Apesar de sua fácil utilização se comparada a linguagens de programação como Java e C++, a assimilação de seus conceitos ainda apresenta acentuada curva de aprendizado para profissionais e pesquisadores não familiarizados com algoritmos e técnicas de programação, conforme constatado em avaliações realizadas por participantes de cursos e apresentações sobre o ambiente *TerraME* (Câmara et al. 2007, Câmara et al. 2008). Entretanto, são esses os profissionais mais interessados no desenvolvimento de modelos para sistemas ambientais e que detêm o conhecimento a respeito do domínio de aplicação dos modelos. Desta maneira, o uso direto e obrigatório de uma linguagem de programação para representação de modelos dinâmicos se apresenta como a principal barreira para a difusão do ambiente *TerraME* e para a ampliação de sua comunidade de usuários.

Para permitir que os usuários do ambiente *TerraME* se concentrem na resolução dos problemas pertencentes ao domínio de aplicação dos modelos e não nos problemas da sua representação computacional, um novo e mais alto nível de abstração é necessário. A representação dos modelos através de componentes gráficos, como diagramas, ao invés de algoritmos, além de aumentar a produtividade dos atuais usuários do *TerraME*, diminuirá a curva de aprendizado de novos usuários.

Este trabalho apresenta o atual estágio de desenvolvimento do *TerraME Graphical Interface for Modeling and Simulation – TerraME GIMS*, uma interface gráfica com o usuário (GUI – *Graphical User Interface*) que permite a descrição e simulação de modelos de sistemas ambientais para a plataforma *TerraME* por meio de metáforas visuais que representam graficamente o modelo.

2. TerraME

O *TerraME* (*TerraLib Modelling Environment*), desenvolvido por Carneiro (2003), é um componente da família de soluções *TerraLib* (Câmara et al. 2000) para a implementação e simulação de modelos ambientais que envolvam a representação explícita do espaço. O *TerraME* provê mecanismos para a representação e simulação de modelos espaciais dinâmicos integrados a um Sistema de Informações Geográficas (SIG). Este ambiente permite a criação de modelos de múltiplas escalas espaciais e temporais.

O *TerraME* foi construído baseado na arquitetura em camadas, onde as camadas inferiores fornecem funcionalidades sobre as quais as camadas superiores são implementadas. Os componentes de sua arquitetura permitem aos usuários experientes a implementação de modelos utilizando diretamente o código fonte do ambiente através da linguagem de programação C++, enquanto aqueles que possuem apenas o conhecimento básico sobre algoritmos e modelagem computacional podem utilizar a linguagem de programação de alto nível *TerraME Modeling Language* – uma extensão da linguagem de programação LUA (Ierusalimschy et al. 1996), que permite a fácil escrita, leitura e alteração dos modelos (Carneiro 2003).

3. Eclipse

Desenvolver um sistema de computação e um modelo computacional para fenômenos sócio-ambientais são atividades essencialmente similares. O uso de um ambiente de desenvolvimento integrado (ou IDE, do inglês *Integrated Development Environment*) cujas características e funcionalidades contribuem para agilizar o processo de desenvolvimento de software é, portanto, também aplicável ao desenvolvimento de modelos para sistemas ambientais. Entretanto desenvolver um ambiente desta natureza possui um custo muito elevado. Uma alternativa à implementação de um ambiente totalmente novo é a utilização de ambientes, ferramentas e *frameworks* já existentes.

O Eclipse (http://www.eclipse.org) é um IDE *open source* que fornece uma plataforma comum para diversos produtos baseados em IDE e facilita a integração destes, servindo de base para a construção de ferramentas e aplicações diversas (Rivieres 2004). Seu grande diferencial é a sua capacidade de integração obtida por meio da sua arquitetura baseada em *plug-ins*. Um *plug-in* é a menor unidade funcional passível de ser desenvolvida e distribuída separadamente. A Plataforma Eclipse é construída num mecanismo de descobrir, integrar, e executar *plug-ins*. Desenvolver uma aplicação sobre a plataforma *Eclipse* permite que ela seja integrada a outras aplicações que também foram escritas sobre esta plataforma. (Eclipse 2006).

A Figura 1 mostra os principais componentes do *Eclipse Software Development Kit* (*Eclipse SDK*) e sua arquitetura baseada em *plug-ins*. O *Eclipse SDK* inclui a Plataforma Eclipse além de duas ferramentas úteis ao desenvolvimento de *plug-ins*: a *Java Development Tools* (JDT), que implementa um ambiente de desenvolvimento Java, e o *Plug-in Development Environment* (PDE) que adiciona ferramentas específicas ao desenvolvimento de *plug-ins* e extensões. Novas aplicações são desenvolvidas estendendo-se o sistema através de *plug-ins* (Eclipse 2006).

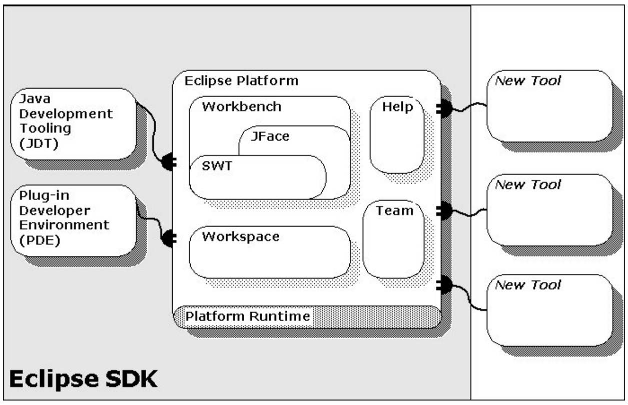


Figura 1. Visão geral do Eclipse SDK (Eclipse 2006)

4. TerraME GIMS

O *TerraME GIMS – TerraME Graphical Interface for Modeling and Simulation*, é uma interface gráfica que permite a construção de modelos dinâmicos espaço-temporais para o ambiente *TerraME* através da interação com componentes gráficos. Está sendo desenvolvido sobre a plataforma Eclipse, implementado e distribuído portanto como um conjunto de *plug-ins* para o Eclipse.

Em conformidade com as plataformas sobre as quais está sendo desenvolvido, *TerraME* e *Eclipse*, o *TerraME GIMS* apresenta uma arquitetura de software em camadas. O *TerraME GIMS* irá compor uma nova camada, entre o *TerraME* e o usuário final, sendo a plataforma Eclipse uma camada intermediária entre o *TerraME* e o *TerraME GIMS*, conforme ilustrado na Figura 2. Desta forma, não há restrição para a criação de modelos diretamente sobre a linguagem *TerraME* quando utiliza-se o *TerraME GIMS*.

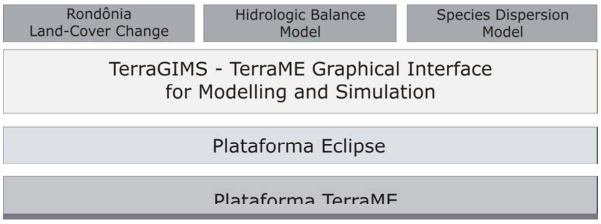


Figura 2. Visão geral da Arquitetura do *TerraME GIMS*

A construção de modelos é feita através do Eclipse e do conjunto de *plug-ins* que constituem o *TerraME GIMS*. A Figura 3 ilustra o ambiente de desenvolvimento dos usuários do *TerraME GIMS*.

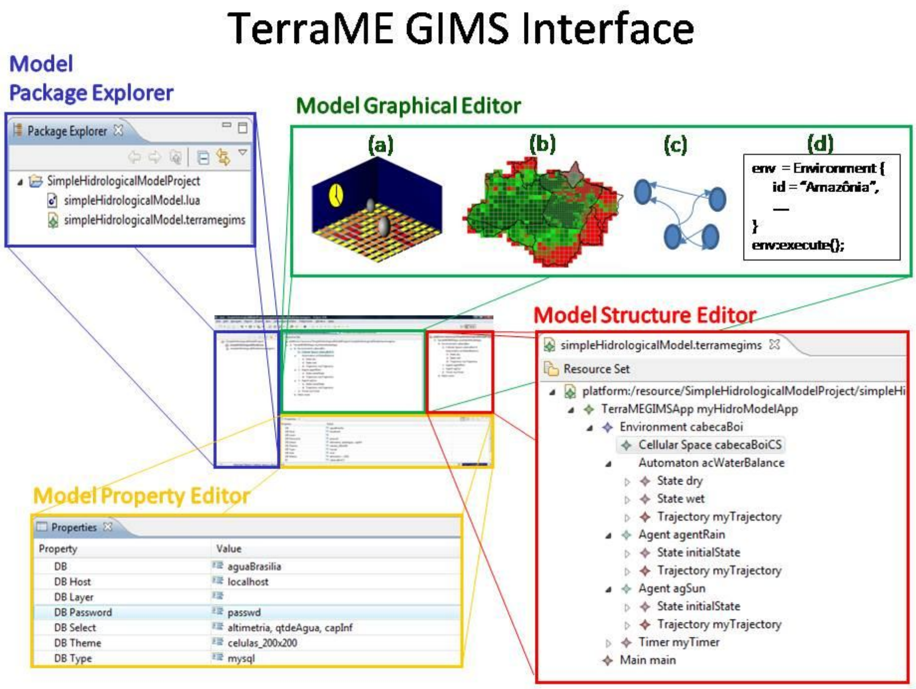


Figura 3. *TerraME GIMS* incorporado ao Eclipse

Através do *Model Package Explorer* os usuários podem acessar os diversos arquivos que fazem parte do projeto. O *Model Structure Editor* permite navegar pelo modelo estruturado hierarquicamente na forma de uma árvore. Uma vez selecionado um elemento do modelo, este poderá ser editado a partir do *Model Graphical Editor* e do *Model Property Editor*.

No *Model Graphical Editor* serão apresentadas ao usuário representações gráficas correspondentes aos diversos elementos do modelo, como por exemplo de um *Environment* em (a), contendo a representação de seus agentes, espaço e tempo; de um *CellularSpace* em (b), exibindo a representação do espaço celular carregado a partir de um banco de dados; de um *Agent* em (c), representado internamente como um autômato, constituído de estados e transições de estados; e o código *TerraME* correspondente em (d), gerado automaticamente em conformidade com o modelo criado pelo usuário a partir da interação com as diversas representações gráficas de seus elementos. As propriedades dos elementos podem ser visualizadas e editadas a partir do *Model Property Editor*, por meio de *widgets* (componentes gráficos de interação com o usuário) como campos de texto e caixas de seleção.

5. Resultados

O trabalho apresenta como resultados parciais de desenvolvimento uma versão do *TerraME GIMS* que permite a criação e visualização dos elementos do modelo representados na forma de uma árvore (*Model Structure Editor*), e a edição das propriedades dos elementos do modelo por meio de campos de texto e menus (*Model Property Editor*). Uma vez criado o modelo, o código *TerraME* correspondente é gerado automaticamente.

Um modelo *TerraME GIMS* é estruturado visualmente para os usuários na forma de uma árvore. A estrutura hierárquica na forma de árvore contribui para a percepção dos usuários na construção de modelos aninhados, com múltiplas escalas. A Figura 4 apresenta um exemplo simples de balanço hídrico, que permite ilustrar a visualização hierárquica do modelo, a visualização das propriedades do elemento *CellularSpace* e o código *TerraME* correspondente.

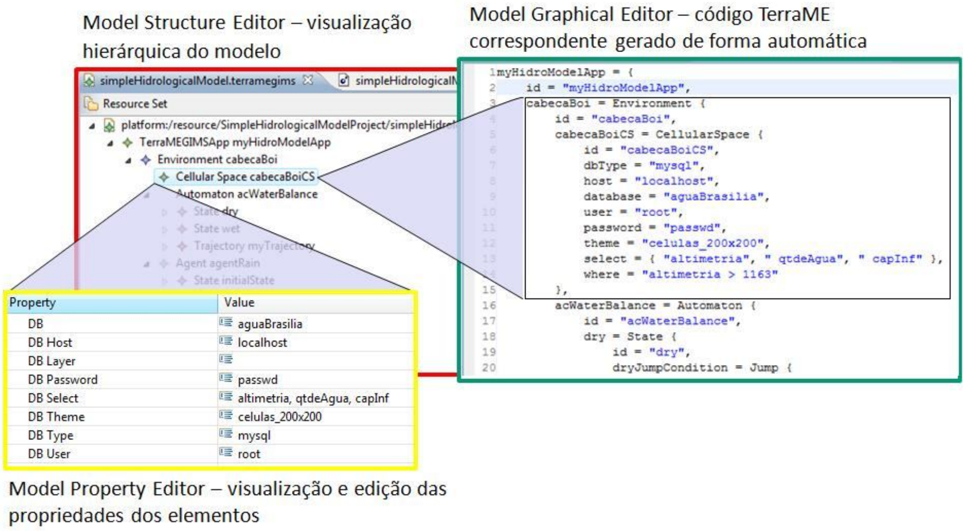


Figura 4. Visualização do hierárquica de um modelo *TerraME GIMS*

Um modelo *TerraME GIMS* é armazenado como um documento XML (http://www.w3c.org/xml), e o código *TerraME* correspondente é gerado (ou atualizado) automaticamente sempre que o modelo *TerraME GIMS* é salvo. A utilização de documentos XML facilita a troca de informações entre diferentes aplicações, contribuindo para integrações futuras de outras ferramentas ao ambiente.

6. Conclusões

A partir dos resultados parciais obtidos com o desenvolvimento do *TerraME GIMS* já é possível vislumbrar seu grande potencial em ampliar a comunidade de usuários do *TerraME*, aumentando a produtividade dos atuais usuários e reduzindo a curva de aprendizado de novos usuários.

Além disto, a partir da utilização do Eclipse como plataforma base de desenvolvimento, uma série de recursos já existentes na plataforma, tais como facilidades para o desenvolvimento colaborativo e controle de versão, poderá ser utilizada na construção de modelos por equipes multidisciplinares e dispersas geograficamente. A arquitetura baseada em *plug-ins* do Eclipse também contribui para futuras extensões ao *TerraME GIMS* e sua reutilização.

As próximas etapas do trabalho incluem incorporar ao *Model Graphical Editor* um editor para construção de autômatos celulares e agentes a partir de diagramas de autômatos (contendo estados e transições de estados), o desenvolvimento de modelos baseados em agentes, e avaliação e testes da interface junto a usuários da aplicação.

6. Agradecimentos

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

References

Carneiro, Tiago G. S. (2003) “Nested-CA: a foundation for multiscale modeling of land use and land change”. São José dos Campos: INPE, 2003.

Câmara, Gilberto. Carneiro, Tiago G. S., Bevilacqua, Luís (2007). I Curso de Verão Geoma 2007 – Modelagem Dinâmica com *TerraME*. http://www.dpi.inpe.br/geoma/curso_verao/index.php. Acessado em Agosto de 2008.

Câmara, Gilberto. Carneiro, Aguiar, Ana P. D., Tiago G. S., Neto, Pedro R., Feitosa, Flávia. (2008). Course on Environmental Modeling – LUCC. http://www.dpi.inpe.br/cursos/environmental_modelling/index.html.

Câmara, G., R. Souza, B. Pedrosa, L. Vinhas, A. Monteiro, J. Paiva, M. Carvalho and M. Gattass (2000). TerraLib: Technology in Support of GIS Innovation. II Workshop Brasileiro de Geoinformática, GeoInfo2000, São Paulo.

Ierusalimschy, R., Figueiredo, L. H., Celes, W. (1996) "Lua - an extensible extension language", Software: Practice & Experience 26 #6. 635-652.

Rivieres, J., Wiegand, J. (2004) Eclipse: A platform for integrating development tools. IBM Systems Journal, 43(2), 2004.

Eclipse (2006) Eclipse Technical Platform Overview. http://www.eclipse.org/articles/Whitepaper-Platform-3.1/eclipse-platform-whitepaper.html. Acessado em Agosto de 2008.

A Progressive Vector Map Browser

José Augusto S. Ramos[1], **Claudio Esperança**[2], **Esteban Walter G. Clua** [1]

[1] Instituto de Computação – Universidade Federal Fluminense (UFF)
Niterói, RJ – Brazil

[2]Programa de Engenharia de Sistemas – COPPE – Universidade Federal do Rio de Janeiro (UFRJ)
Rio de Janeiro, RJ – Brazil

ja_sapienza@yahoo.com.br, esperanc@cos.ufrj.br, esteban@ic.uff.br

Abstract. *With the increasing popularity of web-based map browsers, remotely obtaining a high quality depiction of cartographic information has become commonplace. Most web mapping systems, however, rely on high-capacity servers transmitting pre-rendered tiled maps in raster format. That approach is capable of producing good quality renderings on the client side while using limited bandwidth and exploiting the browser's image cache. These goals are harder to achieve for maps in vector format. In this work, we present an alternative client-server architecture capable of progressively transmitting vector maps in levels-of-detail (LOD) by using techniques such as polygonal line simplification, spatial data structures and, most importantly, a customized memory management algorithm. A multiplatform implementation of this system is described, as well as several performance and image quality measurements taken with it.*

1. Introduction and related work

The increasing popularity of web-based systems and services for delivering maps can be regarded as one of the most important developments in the advancement of cartography. Several aspects of these systems have merited investigation in recent years, such as the improving reliability of the Internet and web server infrastructure, ascertaining the quality and fidelity of the served data, coping with privacy and security issues, maximizing the use of screen space, and making rational use of the available bandwidth [Burghardt et al. 2005].

One important design decision when building a web mapping application is the choice between raster and vector formats. Some of the most popular systems, such as Google Maps [Google, Inc. 2008] or Yahoo Maps [Yahoo! Inc. 2008], are mostly raster-based, which means that they serve pre-rendered digital images, that is, maps are first rendered in several resolutions in the server, cut into blocks (called tiles), and then sent to clients based on the desired area and zoom level. The key reasons for using rasters are: (1) all work spent in authoring a good quality representation can be done on the server, with the merely composing a big picture from several small image tiles; (2) the transmission of raster data of fixed size uses limited bandwidth, and (3) web browsers already manage image caches and, thus, little or no memory management is needed on the client side.

Several web mapping systems have also been proposed which use vector data as well. Perhaps the most widely used system is the MapServer open source project [University of Minnesota 2008]. The technology for serving maps in vectorial form has been intensely researched (see [Kraak and Brown 2001] for a survey). The advent of the SVG (Scalable Vector Graphics) has further propelled these initiatives and general information and software for deploying such systems is abundant (see [Carto:net - cartographers on the net 2008], for instance). An adequate solution for the problem clearly depends on the use of techniques for hierachically organizing vector data according to its visual importance, obtaining what is usually known as level-of-detail data structures [Davis 2008]. At the same time, some sort of spatial indexing is usually required for quickly retrieving data which overlap the region of interest [Gaede and Günther 1998]. Several authors have also investigated suitable techniques for memory caching of spatial data with and without prefetching [Stroe et al. 2000, Doshi et al. 2003], as well as methods appropriate for handling multi-resolution vector data [Chim et al. 1998].

One less studied aspect of web cartography is the relation between the level-of-detail of the data being served, the use of bandwidth and client memory management, specially for vector-based software. In particular, most systems assume that all clients have the same (small) amount of memory at their disposal and, as a consequence, statically link the level-of-detail of the served map to the size of the area being viewed.

In this paper we describe data structures and algorithms which make it possible to remotely deliver and present high-quality vector maps in a progressive manner, making efficient use of the available bandwidth, and adapted to the memory profile of any given client without encumbering the communication protocol with information about client memory state. In particular, although the server receives from the client only information pertaining to the area being viewed, it is able to guess and progressively transmit only needed data.

2. Overall system architecture

According to [McMaster and Shea 1992], around 80% of the total data in vector geographical databases are polygonal lines. This statement guides the scope of the proposed architecture: (1) only vector maps with open or closed polygonal lines are considered, and (2) the use of network bandwidth is optimized by restricting the transmission of line data with just enough detail for a faithful representation.

Thus, we propose a client-server system for serving map data containing a program (the server) capable of directly accessing all polygonal lines of a given map – from a database, for instance – and progressively sending it to interactive visualization applications (the clients). Clients have limited memory capacity and thus store only enough line data so as to present a good depiction of the map within a visualization window. Each time a user changes this window, the contents of the client memory must be updated by requesting relevant data from the server and discarding unneeded information.

The system preprocesses all polygonal lines comprising the map into two hierarchical data structures, which can be quickly traversed in order to obtain the needed information. The two structures used are: (1) a spatial index, which is needed to prune out polygonal lines which do not contribute to the current viewing window, and (2) a structure for organizing the vertices of each polygonal line in order of visual importance – the so-called level-of-detail (LOD) data structure. It should also be mentioned that the present architecture does not handle polygonal line importance classification, i.e., it is considered that all polygonal lines intersecting a given window need to be drawn at some level of detail. Although map visualization applications typically provide some way of establishing which lines can be left out when rendering the map at certain zoom levels, we do not concern ourselves with this feature in this paper.

Both server and client process the viewing window change in a similar manner. The client only needs to inform the server of the new viewing window in order to receive the needed data not yet stored in its memory. This requires that the server is kept aware of the memory state of each client: if there are n lines in a map, the server maintains for each client an array of n integers which map each line to the level of detail in which it is represented in the client's memory. Whenever new vertices need to be sent from server to client, this transmission is broken into blocks of limited size. In other words, the architecture supports progressive transmission of detail information so that the visual quality of the client images improve over time, at a rate that depends solely on the available bandwidth.

3. Server-side preprocessing

For each map being served, their polygonal lines must be submitted to a preprocessing step in order to (1) build a hierarchical level-of-detail data structure for their vertices, and (2) build a spatial index used to support window queries.

3.1. Level-of-detail data structure

We employ the well-known Douglas-Peucker (DP) line smoothing algorithm [Douglas and Peucker 1973]. Although it is not a particularly fast algorithm, running in $O(n \log n)$ at best and $O(n^2)$ in the worst case [Hershberger and Snoeyink 1992], it is ranked as the best when it comes to preserving the shape of the original line [McMaster 1987]. Furthermore, it produces a hierarchical representation which can be used for level-of-detail processing.

The DP algorithm recursively subdivides a polygonal line by selecting the vertex at greatest distance from the line segment defined by the first and last point. Figure 1 illustrates this process for an example line. Observe that the subdivision process can be

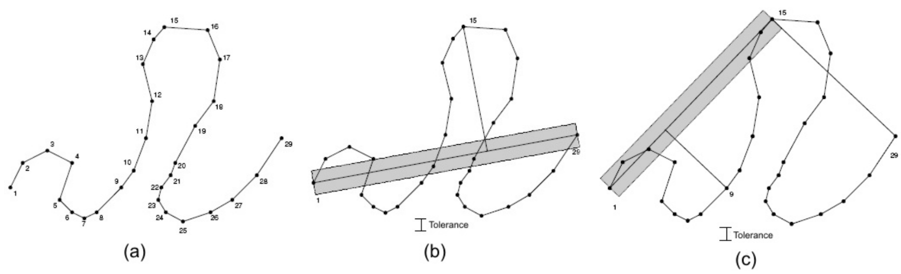


Figure 1. Douglas-Peucker line simplification (adapted from [Casanova et al. 2005] (a) A polygonal line with 29 vertices; (b) vertex 15 is furthest from line segment 1–29, (c) polygonal lines 1–15 and 15–29 are processed recursively.

registered in a binary tree, where each node N corresponds to a subdivision vertex and the corresponding distance d_N, whereas its left and right sons correspond to sub-trees for the left and right sub-chains. Thus, an approximation within tolerance ϵ for the polygonal line can be extracted by visiting the tree in pre-order and pruning out branches rooted at nodes with distances $d_N < \epsilon$.

In this work, we are interested in obtaining increasingly finer representations for each polygonal line. This can easily be implemented by quantizing tolerance values into an integer range $[1, MaxLOD]$. A coarse representation will thus be assigned to level-of-detail (LOD) 1 by visiting the tree in pre-order for tolerance ϵ_1. Vertices in this representation are then marked with $LOD = 1$. The process is repeated for increasingly smaller tolerances ϵ_i for $i = 2 \cdots MaxLOD$, and in each stage i, non-marked vertices are labeled with the corresponding $LOD = i$ value. An important consideration in this process is that, ideally, the number of vertices marked for each LOD value should be approximately constant, so that transmitting the next finer representation of a given line (see constant δ in Section 4) can be done in constant time. In our implementation, ϵ_1 is chosen as the distance in world coordinates corresponding to the width of a pixel for a fully zoomed out projection of the map, while successively finer tolerances were estimated by setting $\epsilon_{i+1} = 0.8\epsilon_i$.

3.2. Spatial indexing

In theory, the worst case scenario for vector map browsing consists of setting the viewing window so that all polygonal lines are enclosed in it. In practice, however, users frequently are interested in investigating a small portion of the whole map. It stands to reason, therefore, that some spatial indexing method be used for selecting polygonal lines intersecting any given query window.

Although the present work does not focus on the issue of efficient spatial indexing, we surveyed several works in the field (see [Samet 2006] for a comprehensive compilation) and chose the relatively simple Expanded MX-CIF Quadtree [Abel and Smith 1984] data structure for speeding up window queries. This is a data structure for rectangles which, in the context of this work, correspond to each polygonal line minimum enclosing bounding box. Each rectangle is represented in the data structure by a collection of enclosing quadtree blocks. In our implementation, this collection contains a maximum of four blocks, although other amounts might also be possible. The four blocks are obtained by determining the minimum enclosing quadtree block, say B, for each rectangle, say R, and then splitting B once to obtain quadtree blocks B_i ($i \in \{$`NW, NE, SW, SE`$\}$) such that R_i is the portion of R, if any, that is contained in B_i. Next, for each B_i we find the minimum enclosing quadtree block, say D_i, that contains R_i. Now, each rectangle is represented by the set of blocks consisting of D_i (refer to Figure 2 for an example). Window queries can be easily computed by means of a recursive descent algorithm on the tree, i.e., start with the root and recursively visit sons if their quadrants intersect the given window.

4. Memory Management

In the context of a client-server system, the issue of memory management should be governed by the following considerations.

Memory capacity: It is assumed that the client memory is bounded by some given constant. At any one time, the client has its memory partially occupied with a subset of

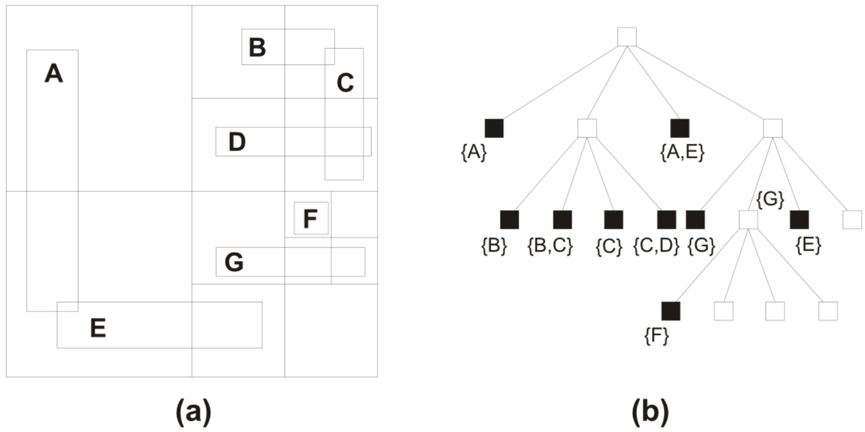


Figure 2. Example MX-CIF Quadtree (b) and the block decomposition induced by it for the rectangles in (a) (adapted from [Samet 2006]).

the map's polygonal lines at some level-of-detail. When the user changes the viewing window, i.e., performs a zooming or panning operation, the memory contents should be altered if it does not contain a "good" representation of the map as revealed by the newly defined window.

Memory control protocol: When requesting new data from the server, some agreement must be reached on what data is needed. In other words, the server must either be told, or must already know what data to transmit to the client in response to a user action. Thus, there are two general approaches for the control protocol: (1) the client requests the needed data items, meaning that the server does not know the client memory's contents, or (2) the server is aware of the memory management operations performed by the client by simulating the same operations as stipulated by a fixed algorithm. Clearly, the former approach uses up more bandwidth than the latter. On the other hand, CPU usage could be greatly increased if the server is to reproduce operations of all clients. In this work, we adopt the second strategy, where the increase in time complexity is alleviated by employing a relatively simple memory management rationale which can be executed in tandem by both server and client.

In order to describe our approach, let us first define a few terms. Let

- M be the maximum client memory size;
- m be the amount of data that can be transmitted from the server to the client in one transaction, i.e., in time for the client displaying the next frame;
- $S = \{L_i\}$ be the current resident set, i.e., the set of polygonal lines L_i that intercept the current viewing window W;
- W be the current viewing window;
- $LOD(L)$ be an integer in $[0, MAXLOD(L)]$ representing the level-of-detail of polygonal line L. It is assumed that if $L \notin S$, then $LOD(L) = 0$;
- $BestLOD(L, W)$ be an estimate for the "best" level-of-detail for exhibiting a polygonal line L in viewing window W. Clearly, this function should return 0 if L is not visible within W. Otherwise, it should return a LOD which is adequate

for the current screen resolution. For instance, it should be just fine enough for mapping two consecutive vertices to different pixels; and

- δ be an estimate of how much memory associated with increasing or decreasing a level-of-detail step for any given polygonal line L. In other words, on average, a polygonal line L should occupy approximately $\delta \times LOD(L)$ memory.

Then, adjusting the items resident in memory requires two classes of operations: operations that increase and operations that decrease the use of memory. Operations in the first class cause data to be transferred from the server to the client. We distinguish two of these:

1. *IncreaseLOD*(L) increases the level-of-detail for polygonal line L. This means that if $LOD(L) = k > 0$, then after its execution $LOD(L) = k + 1$.
2. *Load*(L) brings polygonal line L from the server in its coarsest form. As a precondition, $LOD(L) = 0$ and after its execution, $LOD(L) = 1$.

The second class corresponds to operations which cause data to be thrown away from client memory. Observe that these operations do not cause any network traffic between server and client. We also define two operations of this class:

1. *DecreaseLOD*(L) decreases the level-of-detail for polygonal line L.
2. *Unload*(L) unloads polygonal line L from memory altogether.

Thus, any memory management algorithm will consist of sequentially performing these operations in some order in a timely manner and without infringing memory limits M and m. Our algorithm uses two heaps I and D which hold operations of each of the two classes described above. A crucial consideration is how to define the ordering between operations in each heap. Clearly, operations of type *Load* should have a higher priority than all operations of type *IncreaseLOD*. Similarly, operations of type *DecreaseLOD* should have higher priority than operations of type *Unload*. In our implementation, the ordering between operations *IncreaseLOD* for two lines L_1 and L_2 depend on how distant the $LOD's$ of each line are from their estimated “best”. In other words, we use $|BestLOD(L) - LOD(L)|$ as a priority measure. The priority between operations *DecreaseLOD* is defined in a similar way. Algorithm 1 DefineOperations describes how the two heaps are created.

Once the operation heaps are known, client and server process them in parallel. Operations are executed subject to the memory and bandwidth restrictions discussed above. Algorithm 2 ExecuteOperations summarizes the rationale for operation execution. It is important to realize that executing an operation has different meanings for client and server. For instance, executing an *IncreaseLOD* operation in the client entails receiving line detail from the server and updating the geometry for that line, while for the server it means merely sending the additional vertices. Similarly, while *DecreaseLOD* entails updating the polygonal line data structure for the client, the server needs only to take note that the corresponding memory was freed in the client. A limitation of Algorithm 2 is related to the fact that δ is merely an estimate of the amount of memory associated with decreasing or increasing the LOD of any given line. This may lead to $|S|$, the amount of memory used for the polygonal data, eventually exceeding M if the newly received LOD data is bigger that δ. This, in general, is not a problem since the overflow should be small on average. In any case, the restriction can easily be lifted by assigning to δ a sufficiently large value.

Algorithm 1: DefineOperations

Input: W_{new}: the new window set by the user
Output: I and D: heaps containing operations which cause memory increase/decrease

```
begin
  I ← ∅
  D ← ∅
  S' ← set of lines which intersect W_new
  for L ∈ S' ∪ S do
    if L ∉ S then
      Enqueue [Load, L] in I
    if L ∉ S' then
      Enqueue [Unload, L] in D
    if LOD(L) < BestLOD(L, W_new) then
      for i ← LOD(L) + 1 to BestLOD(L, W_new) do
        Enqueue [IncreaseLOD, L] in I
    else if LOD(L) > BestLOD(L, W_new) then
      for i ← BestLOD(L, W_new) + 1 to LOD(L) do
        Enqueue [DecreaseLOD, L] in D
end
```

Note that the scheme described above is easily adapted to support progressive transmission. Suppose that Algorithm 2 terminates with a non-empty heap I. Then, if the viewing window for the next frame is unchanged, there is no need to run Algorithm 1 again, and the next execution of Algorithm 2 will further process heaps I and D, thus providing an increasingly finer rendering of the map, as illustrated in Figure 7.

5. Implementation and Results

A prototype implementation of the framework described in this paper was built and used to conduct several experiments in order to assess the validity of our proposal.

The development was supported by the following tools: user interfaces were built with version 4.4.1 of the multi-platform *Qt* [TrollTech 2008] library and the *Shapelib* library v. 1.2 was used for reading *Shapefiles* [MapTools 2008]. The pre-processor was written in C++ and compiled using version 4.1 of the gcc compiler [Free Software Foundation Inc. 2008]. The client and server programs which implement the algorithms described above were written in Python (version 2.5.3) with the Qt library accessed by means of the the PyQt wrapper [Riverbank 2008] version 4.4.3. Communication between clients and server use the XML-RPC specification, a protocol for Remote Procedural Call (RPC) coded in XML [UserLand Software, Inc. 2008]. It is important to remark that all of these tools are Open Source and, thus, freely available.

The pre-processing described in Section 4, was carried out with a dedicated program which (1) reads polygonal map data in *Shapefile* format, (2) executes of the Douglas-Peucker algorithm and computes the level-of-detail hierarchy for each polygonal, (3) cre-

Algorithm 2: ExecuteOperations

Input: I and D: heaps containing memory management operations

begin

$t \leftarrow 0$

while $I \neq \emptyset$ *and* $t < m$ **do**

if $|S| + \delta > M$ **then**

$[op, L] \leftarrow$ Dequeue from D

execute $op(L)$

else

$[op, L] \leftarrow$ Dequeue from I

execute $op(L)$

$t \leftarrow t + \delta$

end

Table 1. Sequence of map browsing operations used in the experiments.

Frame	0	8	12	14	16	19	21	23	25	27	31	36	39	41	43	45
Op.	ZF	Z+	P	Z+	Z+	Z+	P	P	P	Z+	Z+	Z-	Z-	Z-	Z-	ZF

ates an extended MX CIF Quadtree for supporting window queries, and (4) saves the important information into a structured XML (eXtensible Markup Language) file which is used as input for the server program.

The system deployment is straightforward, requiring only that a server process is started in some computer and one or more client processes in the same or some other machine connected by a TCP/IP network. When initialized, the server will load the XML generated in the preprocessing stage. When a client connects to the server, it informs its cache memory size and transmission block size, i.e., constants M and m discussed in Section 4. The server then sends a reply message containing the coarsest possible map representation and a compressed representation of the MX-CIF data structure. After this initialization stage, the communication protocol proceeds as described in the above sections.

Experiments were conducted using a test map with the state limits of Brazil in scale 1:1.000.000 produced by the Brazilian Geography and Statistics Institute (IBGE) [IBGE 2008]. This map contains 349 polygonal lines, each with 173 vertices on average. The preprocessing stage produced an Extended MX CIF Quadtree of height 6 and containing 172 leaf nodes, while polygonal lines were split into up to 25 LOD steps.

The maximum client memory size M was set to 32 KBytes, while the maximum block size per frame was set to 2 KBytes. A client was then used for browsing the test map by issuing a fixed sequence of zooming and panning operations. These are shown in Table 5, where ZF, Z+, Z- and P stand for zoom full, zoom in, zoom out and pan, respectively.

The first experiment aimed at measuring the use of client cache memory during the browsing session. The chart in Figure 3 shows that, as expected, the use of memory increases almost linearly and, around frame 19, levels out at the maximum memory capacity.

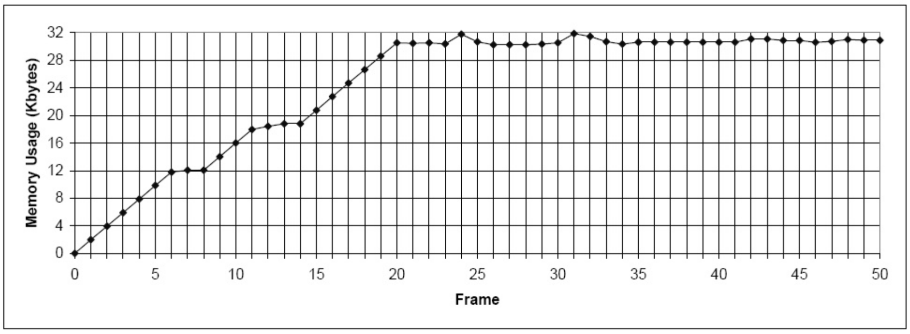


Figure 3. Client cache memory usage per frame.

The second experiment gauged the usage of bandwidth by measuring the sizes of transmitted blocks. These numbers are shown in the chart of Figure 4. As expected, network usage is kept under the imposed limit of 2 KB per frame. Observe that successive frames with high bandwith usage – but without intervening browsing operations – correspond to progressive transmission and rendering of map data. We also note that zooming in and zooming to the full map generate more network traffic than panning or zooming out.

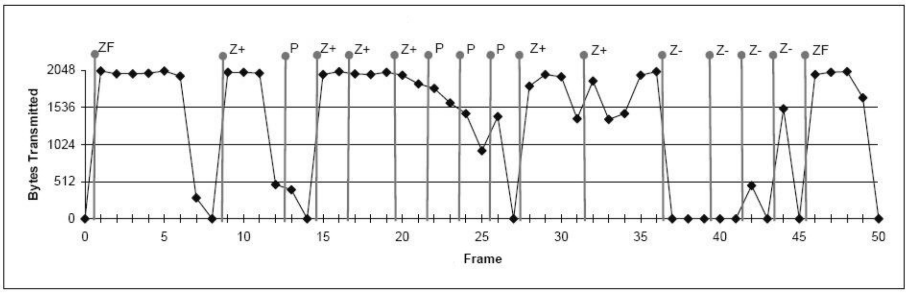


Figure 4. Bytes transmitted per frame.

It is also useful to observe the frequency of cache memory management operations as a function of the frame number. The chart in Figure 5 plots two lines, one depicting the cache inclusion operations, i.e., those that generate network traffic, and another depicting the cache exclusion operations. Note that no exclude operations are required before frame 19, as there is still enough room for storing polygonal data. After this point we observe that the number of include and exclude operations increase and decrease in sync. This is reasonable since exclude operations are required to make room for include operations. Another important observation is that the number of operations does not necessarily follow the pattern of bandwidth usage. This can be attributed to the fact that the amount of data for each LOD step is not constant.

Finally, it was sought some way for measuring the picture quality observed in the client as a function of time (frame). For this purpose, we considered that a given polygonal line L present in window W is rendered perfectly if it is represented in cache

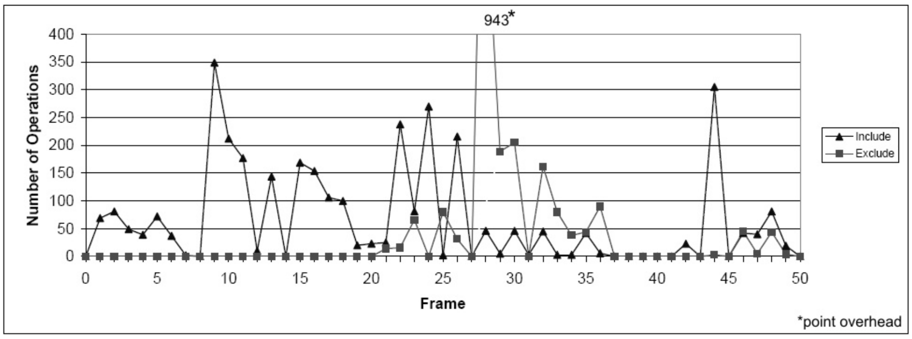


Figure 5. Cache memory management operations per frame.

memory with LOD $BestLOD(L, W)$ or greater. Thus, a percentage measure of image quality Q may be estimated by

$$Q = \frac{100}{|R|} \times \sum_{L \in R} \frac{\min(LOD(L), BestLOD(L, W))}{BestLOD(L, W)},$$

where R is the set of lines intersecting W. A plot of this quality measure is shown in Figure 6. It is possible to observe that after a few frames the system achieves maximum quality and never falls significantly below 80%. Obviously, this threshold is dependent on the relationship between the total map size and M, the cache memory size. Similarly, the latency observed for reaching maximum quality depends on the allowed block size m.

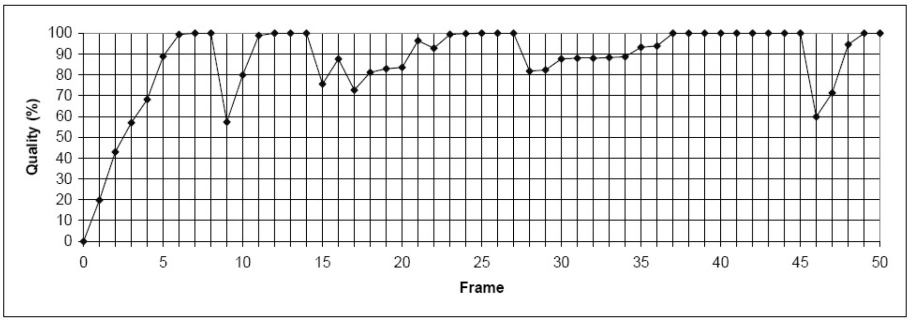


Figure 6. Image quality per frame.

As an example of quality increasing over time is shown in Figure 7. The user performs a zoom operation in a window displaying a map with all states of Brazil, causing the system to perform a progressive transmission and rendering until it reaches 100% of image quality.

6. Conclusions and suggestions for future work

The client-server framework for remotely displaying vector maps described in this work was designed to achieve several goals: be simple, scalable, make predictable use of network bandwidth and support progressive transmission and rendering. The prototype implementation and admittedly limited experimental evidence seem to indicate that these objectives were largely met. A continuation of this work would necessarily start by a more thorough experimentation, with the use of different data sets and other values for

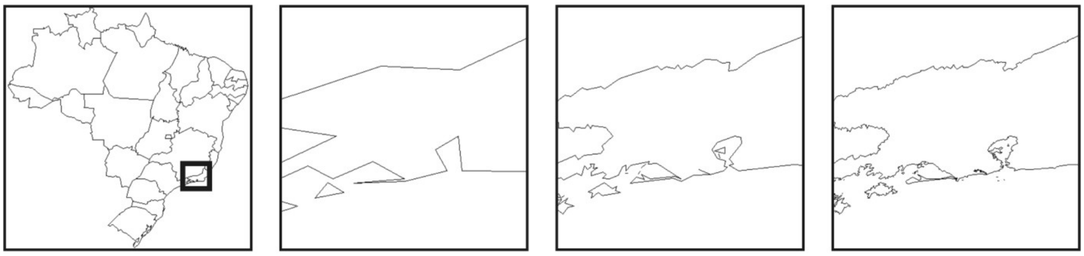


Figure 7. A zoom-in operation performed on the black rectangle (a) causes progressive transmission and rendering (b), (c) until the system achieves maximum image quality (d).

constants M and m. Ideally, a performance evaluation should also be attempted in order to evaluate the scalability of the server.

Clearly, a production system would require the addition of several improvements such as visual importance classification and a more fine-grained processing of polygonal data so that a given polyline could be stored with varying levels of detail in client memory. A complete system would probably also include the ability to serve raster data.

References

Abel, D. J. and Smith, J. L. (1984). A data structure and query algorithm for a database of areal entities. *Australian Computer Journal*, 16(4):147–154.

Burghardt, D., Neun, M., and Weibel, R. (2005). Generalization services on the web – a classification and an initial prototype implementation. *Proceedings of the American Congress on Surveying and Mapping – Auto-Carto*.

Carto:net - cartographers on the net (2008). Svg, scalable vector graphics: tutorials, examples, widgets and libraries. `http://www.carto.net`.

Casanova, M., Câmara, G., Davis, C., Vinhas, L., and Queiroz, G. (2005). *Bancos de dados geográficos*. Editora Mundo GEO, Curitiba.

Chim, J. H. P., Green, M., Lau, R. W. H., Leong, H. V., and Si, A. (1998). On caching and prefetching of virtual objects in distributed virtual environments. In *MULTIMEDIA '98: Proceedings of the sixth ACM international conference on Multimedia*, pages 171–180, New York, NY, USA. ACM.

Davis, C. (2008). Geometria computacional para sistemas de informação geográfica. `http://www.dpi.inpe.br/~gilberto/livro/geocomp/`.

Doshi, P. R., Rundensteiner, E. A., and Ward, M. O. (2003). Prefetching for visual data exploration. In *DASFAA '03: Proceedings of the Eighth International Conference on Database Systems for Advanced Applications*, page 195, Washington, DC, USA. IEEE Computer Society.

Douglas, D. H. and Peucker, T. K. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 2(10):112–122.

Free Software Foundation Inc. (2008). GCC, the GNU compiler collection. http://gcc.gnu.org/.

Gaede, V. and Günther, O. (1998). Multidimensional access methods. *ACM Comput. Surv.*, 30(2):170–231.

Google, Inc. (2008). Google maps. http://maps.google.com.

Hershberger, J. and Snoeyink, J. (1992). Speeding up the Douglas-Peucker line-simplification algorithm. In *Proc. 5th Intl. Symp. on Spatial Data Handling*, pages 134–143.

IBGE (2008). IBGE - Instituto Brasileiro de Geografia e Estatística. http://www.ibge.gov.br.

Kraak, M.-J. and Brown, A. (2001). *Web Cartography – Developments and prospects*. Taylor & Francis, New York.

MapTools (2008). Shapefile C library v1.2. http://shapelib.maptools.org/.

McMaster, R. B. (1987). Automated line generalization. *Cartographica*, 24(2):74–111.

McMaster, R. B. and Shea, K. S. (1992). *Generalization in digital cartography*. Association of American Geographers, Washington, D.C.

Riverbank (2008). Pyqt4 download. http://www.riverbankcomputing.co.uk/software/pyqt/download.

Samet, H. (2006). *Foundations of Multidimensional and Metric Data Structures*. Morgan-Kaufman, San Francisco.

Stroe, I. D., Rundensteiner, E. A., and Ward, M. O. (2000). Scalable visual hierarchy exploration. In *DEXA '00: Proceedings of the 11th International Conference on Database and Expert Systems Applications*, pages 784–793, London, UK. Springer-Verlag.

TrollTech (2008). Qt cross-platform application framework. http://trolltech.com/products/qt/.

University of Minnesota (2008). UMN MapServer. http://mapserver.gis.umn.edu.

UserLand Software, Inc. (2008). Xml-rpc homepage. http://www.xmlrpc.com.

Yahoo! Inc. (2008). Yahoo! maps. http://maps.yahoo.com.

Um Framework para Coleta e Filtragem de Dados Geográficos Fornecidos Voluntariamente

João Carlos Tavares da Silva[1], Clodoveu Augusto Davis Jr.[2]

[1]Instituto de Informática – Pontifícia Universidade Católica de Minas Gerais

[2]Departamento de Ciência da Computação – Universidade Federal de Minas Gerais

jsilva27@gmail.com, clodoveu@dcc.ufmg.br

Resumo. *Diversas aplicações na Web permitem que usuários contribuam com dados textuais, fotos, opiniões e recomendações. O mesmo tem ocorrido em sites ligados a informação geográfica. A contribuição voluntária gera freqüentes dúvidas quanto à sua confiabilidade e qualidade. No entanto, é grande a importância potencial desse tipo de contribuição para a expansão e o melhoramento das fontes de dados geográficos atualmente disponíveis. Este trabalho apresenta uma iniciativa de criação de um framework para coleta de contribuições voluntárias sobre informação geográfica na Web, que será usado para pesquisa sobre aspectos como qualidade, confiabilidade, mecanismos de filtragem e validação, e integração a infra-estruturas de dados espaciais.*

1. Introdução

O compartilhamento de grandes volumes de informação espacial vem se tornando cada vez mais importante. Sistemas de Informação Geográficos (SIG) estão se tornando núcleos de ambientes computacionais que envolvem grandes quantidades de usuários distribuídos em rede [7]. Várias ferramentas passaram a permitir o acesso a grandes volumes de dados no formato de mapas e imagens de satélites, viabilizando o uso de informação geográfica em problemas do cotidiano. Essas iniciativas contribuíram também para desmistificar antigos conceitos a respeito da viabilidade do acesso a SIG através da Internet.

Uma conseqüência da grande oferta de dados espaciais básicos online foi o surgimento de aplicações que permitem que uma pessoa acrescente dados sobre o espaço em que vive [2]. Isto resulta, entre outras coisas, na possibilidade de localizar informações de interesse pessoal e anotar melhoramentos e correções que podem ser úteis em futuras revisões de mapas. A proliferação destes aplicativos, baseados no conceito da Web 2.0, abre as portas para a inovação. Milhões de usuários não são mais apenas vistos como consumidores, mas também são colaboradores e criadores de informação [4].

Um problema que surge a partir dessa perspectiva é filtrar as contribuições legítimas, separando-as de conteúdo falso ou mal-intencionado. Existem algumas iniciativas a esse respeito em aplicativos típicos da Web 2.0, mas ainda não há mecanismos consagrados voltados para informação geográfica. Em vista disso, este trabalho propõe o desenvolvimento de um *framework* para coleta de contribuições voluntárias a dados geográficos na Web. Como forma de avaliar a confiabilidade das contribuições, o aplicativo prevê a incorporação, para experimentação e testes, de diferentes mecanismos de filtragem do conteúdo, voltados para estabelecer a confiabilidade dos dados e de seus fornecedores. O objetivo é determinar quais são os mecanismos de controle que se mostram mais eficazes para essa finalidade. Além disso, como se trata de um *framework*, o aplicativo desenvolvido permite que Web sites voltados para contribuição voluntária sejam criados e publicados por pessoas ou organizações interessadas.

Este artigo está organizado da seguinte forma. A seção 2 descreve iniciativas atuais em torno da contribuição de usuários para a formação de conteúdo na Web. A seção 3 apresenta o framework proposto. Na seção 4 são feitas considerações finais e apresentado o que se planeja para a condução das próximas etapas do trabalho.

2. Contribuições voluntárias para dados geográficos

Tem crescido rapidamente o interesse das pessoas por recursos e ferramentas que permitem que conteúdo produzido por indivíduos e grupos seja publicado na Internet, obtendo imediatamente alcance mundial. São inúmeras as iniciativas de criação de blogs, publicação de vídeos domésticos, registro de comentários e opiniões sobre produtos, serviços, livros e filmes, entre muitos outros. O crescimento explosivo de sites que permitem algum tipo de participação dinâmica dos usuários é indicação clara dessa tendência. Como exemplo, pode-se citar sites como YouTube, Orkut, Flickr, del.icio.us, Blogger e muitos outros.

Um exemplo fundamental das possibilidades da criação de conteúdo por parte de usuários é a Wikipedia. Os mecanismos de criação e manutenção da Wikipedia a colocam no centro de uma grande polêmica, envolvendo questões como qualidade, confiabilidade, completeza e atualidade dos artigos, em comparação com fontes tradicionais de informação. Mas é inegável que o modelo *wiki* de participação voluntária, irrestrita e, na maioria das vezes, anônima, pode levar a resultados que estão se tornando impraticáveis para organizações tradicionais. Basta tentar imaginar o custo e o tamanho do esforço necessário para criar algo como a Wikipedia sob a égide de uma empresa comercial, que teria que pagar pelas contribuições e cobrar pelo acesso ao resultado.

A analogia da Wikipedia com uma enciclopédia convencional, como a Encyclopaedia Britannica, é inevitável. Embora a Britannica tenha desenvolvido uma versão online, o conteúdo integral é reservado para assinantes; embora disponha de uma interface para que leitores cadastrados contribuam com sugestões, não é possível saber se a sugestão foi ou será aceita, e se alguma modificação ou atualização resultou da contribuição do leitor. Enquanto a Britannica é autoritativa, em muitos artigos da Wikipedia há um aviso informando o leitor de que há polêmica sobre o conteúdo apresentado. Mesmo assim, existe ainda muita discussão sobre a confiabilidade de partes do conteúdo da Wikipedia [15].

Os mecanismos que levam as pessoas a dispender tempo e esforço para produzir de graça algo que pode ou não ser usado por outros ainda não são claros [4, 9]. No entanto, a regra básica dos mecanismos *wiki* (e originária do desenvolvimento colaborativo de software livre) parece ser válida para uma ampla gama de situações: "se houver olhos suficientes, todos os *bugs* são visíveis" [14]. Ou seja, se um conjunto de informações disponível livremente na Web for de interesse para um grande número de pessoas, uma parcela razoável dessas pessoas tende a retribuir o serviço prestado por meio de uma predisposição a colaborar para que o serviço melhore. Tal fenômeno tem sido estudado na área de Comunicação, sendo conhecido como “ação coletiva” [3, 8], e voltado para acervos públicos de informação, chamados de *information commons* [13].

Com relação à informação geográfica, um fenômeno semelhante vem ocorrendo recentemente. Tendo por base acervos de mapas e imagens georreferenciadas disponíveis gratuitamente pela Web (tais como Google Maps, Microsoft Windows Live Local, Yahoo Maps), vários *Web sites* têm sido criados contando com a perspectiva da colaboração do usuário na criação e manutenção de um acervo temático ou complementar de dados geográficos. Por meio desses mecanismos, um usuário qualquer pode tentar reconhecer aspectos da realidade que o cerca e criar uma anotação, comentário, dar um nome, ou conectá-lo a alguma outra fonte de informação complementar. Um exemplo é o site Wikimapia[1], em que usuários delimitam áreas de seu interesse com retângulos ou polígonos e associam a eles descrições e comentários. Muitas vezes, o local indicado recebe também um link para um artigo na Wikipedia, usando o mecanismo de geotags. Outro exemplo é o site OpenStreetMap[2], em que usuários podem editar mapas de ruas a partir de imagens de satélite de alta resolução, dentre as quais as do Yahoo! Maps. Os mapas resultantes podem ser exportados e usados em outras ferramentas de geoprocessamento, estando livres de licenciamento. No Brasil, uma experiência semelhante é o projeto Tracksource[3], iniciado por jipeiros interessados em trocar dados geográficos sobre estradas vicinais e trilhas, coletados por receptores GPS de navegação [11]. O projeto inclui muitos mapas urbanos roteáveis, contendo informação sobre regras de trânsito para planejamento de deslocamentos. Os mapas disponíveis no Tracksource podem ser usados em computadores pessoais ou carregados em receptores GPS para uso móvel em navegação. O Tracksource realiza hoje mais de 50.000 downloads de mapas por mês, totalizando cerca de 400 GB/mês. O site freqüentemente ultrapassa 1.500.000 “hits” mensais, demonstrando o interesse que existe atualmente nesse tipo de informação – bem como as deficiências dos mapas disponíveis para receptores GPS.

Existem precedentes para a contribuição voluntária de dados, inclusive em projetos de pesquisa científica. É conhecida nos Estados Unidos uma iniciativa anual de observação e monitoramento de aves migratórias, em que a posição de bandos de aves em determinado instante é registrada e comunicada aos cientistas aficionados do *hobby* de observação de pássaros. Um exemplo conhecido de contribuição voluntária para mapeamento é o projeto *National Map Corps*, do *United States Geological Survey* (USGS), a agência de mapeamento nacional norte-americana [1, 5]. O USGS iniciou esse projeto na década de 1990, chegando a reunir mais de 3.000 voluntários em 2001. Recentemente, o projeto passou a receber apenas contribuições via Web[4] e que fossem adequadamente acompanhadas de coordenadas obtidas por GPS, mas mesmo assim existe um *backlog* de contribuições que precisam ser tratadas. Na recepção de contribuições via Web, o projeto foi inspirado em uma iniciativa da NASA, denominada criativamente *Clickworkers*[5].

---

[1] http://www.wikimapia.com

[2] http://www.openstreetmap.org

[3] http://www.tracksource.org.br

[4] http://ims.er.usgs.gov/vfs/faces/index.jspx

[5] http://clickworkers.arc.nasa.gov/top

Em resumo, pode-se observar que existem muitas oportunidades para uma avaliação mais científica das iniciativas de participação voluntária na criação e manutenção de acervos de informação geográfica, visando prover melhores serviços no futuro.

3. Framework para coleta e filtragem de contribuições voluntárias

O propósito desta pesquisa é projetar e implementar um *framework* para coleta e dados geográficos fornecidos voluntariamente, através da Web. Por meio do *framework*, uma pessoa ou organização publica na Web um conjunto de dados geográficos, e abre a outros usuários a possibilidade de contribuir com novos dados, propostas de correção ou de atualização. Existe a opção de permitir apenas a carga de novos dados em feições previamente modeladas, ou permitir a criação de temas novos. A Figura 1 apresenta esquematicamente o framework proposto e desenvolvido até o momento.

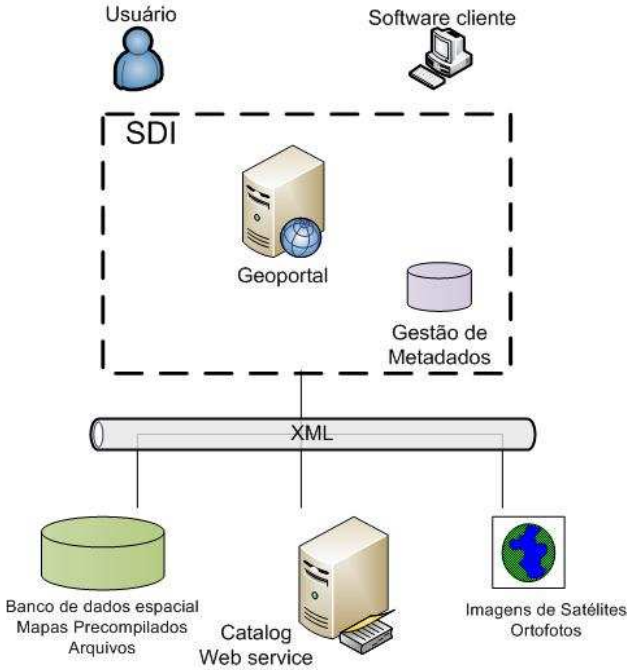


Figura 1 – Diagrama esquemático do framework desenvolvido

A solução, implementada sob um Geoportal, permite ao usuário a customização de sua interface, tornando-a aderente às particularidades e demandas de seu interesse, através de funções que possibilitam a criação de feições personalizadas, a carga de dados a partir de arquivos georreferenciados e a tematização dos dados (Figura 2). A criação de feições personalizadas torna a solução ajustável às necessidades do usuário, permitindo a inclusão de dados cujos atributos e propriedades são definidos pelos próprios usuários. É possível determinar o nome da feição, seu tipo de geometria, o número de atributos, o tipo e o valor padrão associado a cada atributo, as informações principais para exibição no mapa, a escala de visualização e a simbologia relacionada a cada faixa de escalas. Para cada feição personalizada, o sistema cria uma camada que permite a exibição dos dados correspondentes no mapa. Após a criação de uma feição, os objetos espaciais relacionados a ela podem ser adicionados manualmente ou por carga de dados a partir de um arquivo.

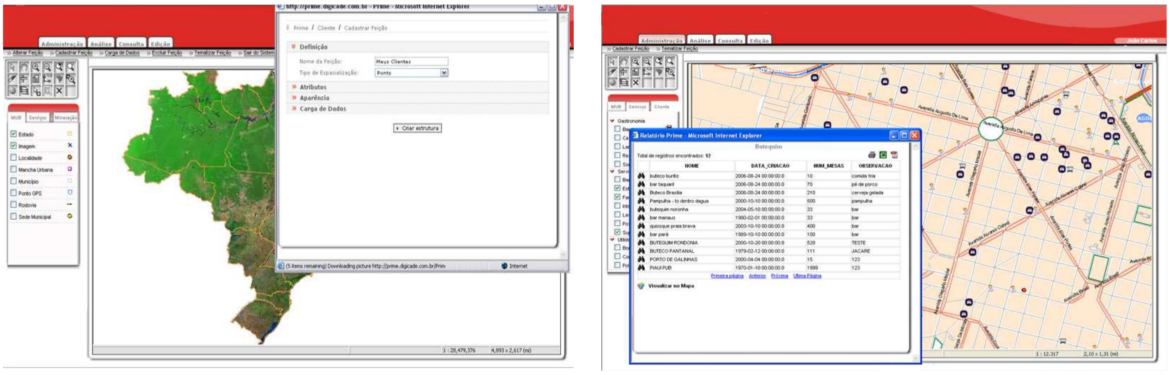


Figura 2 – Interfaces da solução implementada: telas para cadastro e inserção de dados

O *framework* utiliza um gerenciador de bancos de dados geográficos para armazenar o esquema criado pelo usuário para suas feições. No mesmo gerenciador são armazenados os dados de fundo (mapas básicos) e os dados fornecidos pelo usuário. A Figura 3 exibe um fluxograma de funcionamento da aplicação. Criada a estrutura, pode ser liberado o acesso para que outras pessoas possam contribuir, ou ainda continuar o cadastro de novas feições de interesse. O objetivo é permitir o acesso por meio de um visualizador, instalado como componente de um geoportal, e também publicar os dados fornecidos pelo usuário utilizando serviços Web padrão OGC. Com isso, os dados fornecidos pelos voluntários podem se tornar parte de uma infra-estrutura de dados espaciais (IDE) [6].

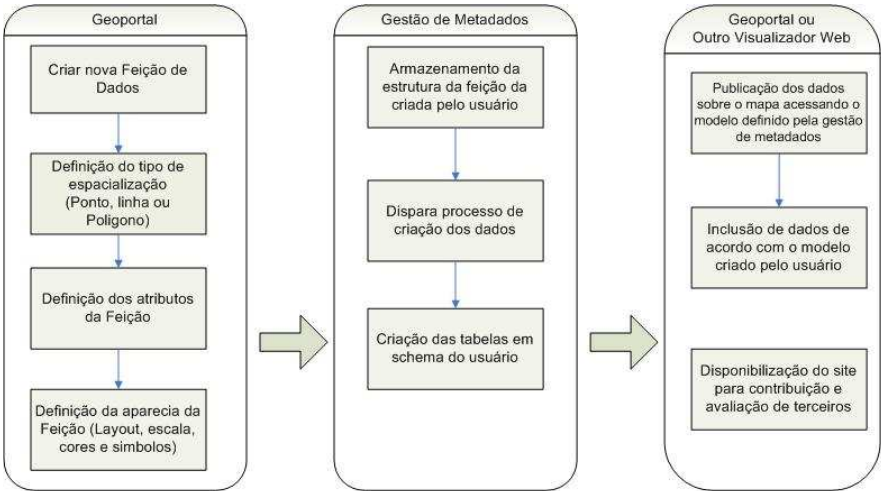


Figura 3 - Fluxograma de funcionamento da solução

4. Considerações finais e trabalhos futuros

Este projeto está em andamento. No estágio atual, usuários serão apresentados ao *framework*, e mecanismos de filtragem estão sendo implementados. A idéia é criar comunidades de usuários que contribuam com dados geográficos sobre um tema de interesse comum, em um ambiente controlado. Será também desenvolvida uma adaptação para que o usuário possa visualizar seus dados fora do ambiente do Geoportal, usando serviços Web OGC. O *framework* incluirá ainda um conjunto de recursos para que o criador de uma camada de dados ou feição possa solicitar, aos próprios voluntários, uma avaliação da qualidade dos dados fornecidos [8, 10]. Diversos mecanismos como esses foram propostos e estão disponíveis para validação de contribuições voluntárias, geográficas ou não, em Web sites do tipo *wiki*. Esses mecanismos serão implementados e integrados ao *framework*. A idéia é permitir que o criador de um *site* ou feição possa especificar o tipo de mecanismo que considera mais interessante, levando em conta a natureza dos dados e da comunidade envolvida no processo. Esse recurso abre diversas possibilidades interessantes para pesquisa, na avaliação comparativa entre a qualidade percebida dos dados fornecidos e a qualidade efetiva, determinada por especialistas. Também permite comparar mecanismos de filtragem de contribuições voluntárias, de modo a apontar os mais eficientes, e pode levar à implementação de sistemas de reputação e recomendação para dados geográficos na Web [12].

5. Referências

1. Bearden, M.J. *The National Map Corps*. in *Workshop on Volunteered Geographic Information*. 2007. Santa Barbara, California, USA.
2. Bell, T., Cope, A.S., and Catt, D. *The third spatial revolution*. in *Workshop on Volunteered Geographic Information*. 2007. Santa Barbara, California, USA.
3. Bimber, B., Flanagin, A.J., and Stohl, C., *Reconceptualizing collective action in the contemporary media environment.* Communication Theory, 2005. **15**(4): p. 365-388.
4. Budhathoki, N.R. *Reconceptualization of user is essential to expand the voluntary creation and supply of spatial information*. in *Workshop on Volunteered Geographic Information*. 2007. Santa Barbara, California, USA.
5. Cowen, D.J. *Why not a geo-wiki corps?* in *Workshop on Volunteered Geographic Information*. 2007. Santa Barbara, California, USA.
6. Davis Jr, C.A. and Alves, L.L. *Local Spatial Data Infrastructures Based on a Service-Oriented Architecture*. in *8th Brazilian Symposium on GeoInformatics (GeoInfo 2005)*. 2005. Campos do Jordão (SP).
7. Davis Jr., C.A. and Fonseca, F.T., *Considerations from the Development of a Local Spatial Data Infrastructure.* Information Technology for Development, 2005. **12**(4): p. 273-290.
8. Frew, J. *Provenance and volunteered geographic information*. in *Workshop on Volunteered Geographic Information*. 2007. Santa Barbara, California, USA.
9. Goodchild, M.F., *Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0.* International Journal of Spatial Data Infrastructures Research, 2007. **2**: p. 24-32.
10. Gouveia, C. *How to make sense of citizen participation in environmental monitoring?* in *Workshop on Volunteered Geograhic Information*. 2007. Santa Barbara, California, USA.
11. Lima, M.F.M., *O Projeto Tracksource*, in *Instituto de Educação Continuada*. 2008, Pontifícia Universidade Católica de Minas Gerais.
12. Maué, P. *Reputation as a tool to ensure validity of VGI*. in *Workshop on Volunteered Geographic Information*. 2007. Santa Barbara, California, USA.
13. Onsrud, H., Câmara, G., Campbell, J., and Chakravarthy, N.S., *Public Commons of Geographic Data: research and development challenges*, in *GIScience 2004 (LNCS 3234)*, Egenhofer, M.J., Freksa, C., and Miller, H.J., Editors. 2004, Springer-Verlag: Berlin/Heidelberg. p. 223-238.
14. Raymond, E. *The Cathedral and the Bazaar*. 1999 [cited; Available from: Available at http://www.tuxedo.org/~esr/writings/cathedral-bazaar/.
15. Waters, N.L., *Why you can't cite Wikipedia in my class.* Communications of the ACM, 2007. **50**(9): p. 15-17.

Detecção Automática de Rotas de Ônibus

Leone Pereira Masiero, Marco A. Casanova, Marcelo T. M. Carvalho

Laboratório de Tecnologia em Computação Gráfica (TeCGraf)
Departamento de Informática
Pontifícia Universidade Católica do Rio de Janeiro – Rio de Janeiro, RJ – Brasil

{leone,casanova,tilio}@tecgraf.puc-rio.br

Abstract. *This paper describes an algorithm for machine learning of routes traveled by an urban bus fleet by analyzing the sequences of points (geo-referenced) sent by devices installed in each bus. The work includes tests with real data and also discusses some practical aspects that can detect the street segments of any errors inherent in the use of GPS.*

Resumo*. Este trabalho descreve um algoritmo para aprendizagem automática de rotas percorridas por uma frota de ônibus urbanos através da análise das seqüências de pontos (geo-referenciados) enviados por dispositivos instalados em cada ônibus. O trabalho inclui testes com dados reais e discute ainda alguns aspectos práticos para detecção de trechos decorrentes dos erros inerentes ao uso de GPS.*

1.Introdução

A correlação de eventos em tempo real [1,6] é uma tecnologia importante no contexto de gerência de redes, detecção de intrusos [2] e processamento de logs gerados por sistemas [3]. No contexto de geoprocessamento dinâmico, Hornsby [5] argumenta que uma possível área de interesse consiste na identificação de rotas percorridas por objetos móveis ou a detecção de eventos de interesse em um grupo de objetos móveis, como um conjunto de navios partindo de um porto devido a uma explosão. O monitoramento de frotas [4] oferece um segundo exemplo de aplicações de geoprocessamento que dependem de processamento de eventos em tempo real.

Este trabalho concentra-se no problema de aprendizagem automática de rotas percorridas por uma frota de ônibus urbanos através da análise das seqüências de pontos (geo-referenciados) enviados por dispositivos instalados em cada ônibus. Este problema é interessante por várias razões. Primeiro, o monitoramento de uma frota de ônibus urbanos naturalmente necessita que as rotas que os ônibus seguem sejam definidas com precisão. Sem isto, é impossível detectar, entre outros pontos, se um ônibus está fora de rota, ou se está trafegando acima do limite de velocidade para o trecho. Porém, a definição manual de cada rota é um processo bastante laborioso, sujeito a variações sazonais, devido a feriados, obras temporárias, entre outros fatores. De fato, a solução adequada para este problema provou ser um fator importante para o sucesso da implementação de uma aplicação de monitoramento de frotas realizada por um dos autores. Por outro lado, pela própria natureza de uma linha urbana de ônibus, os trajetos percorridos sucessivamente pelos ônibus repetem-se dentro de um mesmo período do dia ou dentro de um intervalo de dias específico (como durante o carnaval). Assim, é viável construir um algoritmo que aprenda rotas analisando as sequências de pontos enviados por dispositivos instalados em cada ônibus e acessando o mapeamento das ruas da cidade coberta pela linha de ônibus.

O algoritmo para aprendizagem de rota apresentado neste trabalho é parte de um esforço mais amplo para construir um sistema de monitoramento de frotas de ônibus, em andamento no Laboratório de Tecnologia em Computação Gráfica (TeCGraf) do Departamento de Informática da Pontifícia Universidade Católica do Rio de Janeiro. O sistema está estruturado em módulos separados para recepção, interpretação, controle, correlação, armazenamento persistente e visualização de seqüências de dados enviados por dispositivos instalados nos ônibus.

Este trabalho está organizado da seguinte forma. A seção 2 apresenta alguns conceitos relacionados à detecção de rotas e a organização do banco de dados. A seção 3 apresenta o algoritmo de detecção de rotas. A seção 4 apresenta os testes realizados. Por fim, a seção 5 contém as conclusões do trabalho.

2. Modelo de Rotas

Um *trecho* é uma partição indivisível de uma rua, de modo que seja vedada a qualquer ônibus a possibilidade de percorrer apenas uma fração de um trecho. Desta forma, trechos só podem ser conectados entre si pelas suas extremidades. Entretanto, os trechos podem apresentar forma totalmente arbitrária, sendo assim representados através de polilinhas. Neste trabalho, convencionamos que um trecho pode ser trafegado em ambos os sentidos.

Um *caminho* é uma seqüência $(S_1,...,S_n)$ de trechos, onde cada trecho está associado a uma restrição temporal (intervalo de tempo). Uma *rota* é uma composição $(P_1,...,P_n)$ de caminhos, onde todos os caminhos da rota devem começar no mesmo trecho S_1 e terminar no mesmo trecho S_n. Todos os caminhos a serem percorridos por veículos de de uma mesma rota devem ter no mínimo um trecho intermediário diferente entre todos eles, como é mostrado na figura 1. Uma rota é *bidirecional* quando pode ser percorrida em ambas as direções; caso contrário é *unidirecional*. Uma rota é *circular* quando o primeiro ponto de S_1 coincide com o último ponto de S_n.

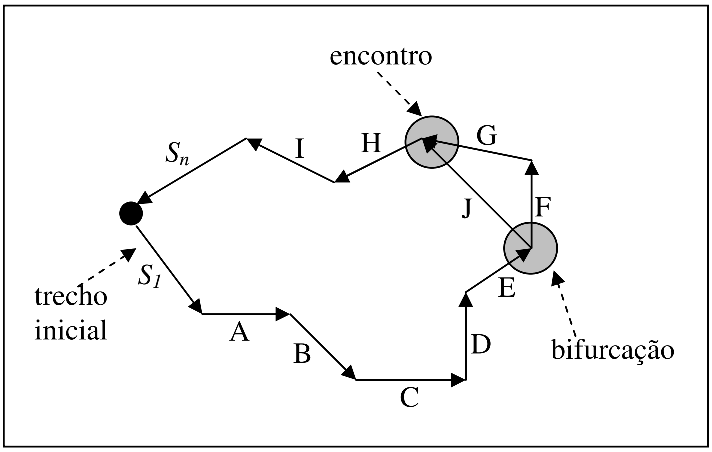


Figura 1 – Rota circular com 2 caminhos.

O modelo do banco de dados deve refletir exatamente estas definições de trecho. Se o mapeamento urbano não considera a noção de trecho, é necessário particionar cada rua em trechos, de forma apropriada. Cada trecho possuirá atributos previamente determinados, como a velocidade máxima permitida. Assim, se há uma escola, por exemplo, onde a velocidade máxima permitida é diferente do resto da rua, deve-se tomar o cuidado para definir o trecho da escola como um trecho separado.

Dada a representação conceitual de uma rota com as definições de trecho e caminho, a representação computacional de uma rota pode ser definida por um grafo direcionado. Cada vértice do grafo contém um conjunto de trechos subsequentes que não tenham entre si bifurcações nem encontros, e um intervalo de tempo que representa a união dos intervalos de tempo dos trechos que pertencem ao vértice. Por exemplo, o vértice do grafo da figura Figura 2 que contém os trechos F e G terá como intervalo de tempo a união dos intervalos de tempo dos trechos F e G. Se o trecho F possuir um intervalo de tempo entre t_1 e t_2 (onde $t_2 > t_1$) e o trecho G possuir um intervalo de tempo entre t_3 e t_4 (onde $t_4 > t_3 > t_2$), a união desses dois intervalos seria um intervalo de tempo entre t_1 e t_4, que determina o intervalo de tempo do vértice do grafo que contém os dois trechos F e G. As arestas do grafo representam as conexões entre os trechos, como pode ser visto na figura 2.

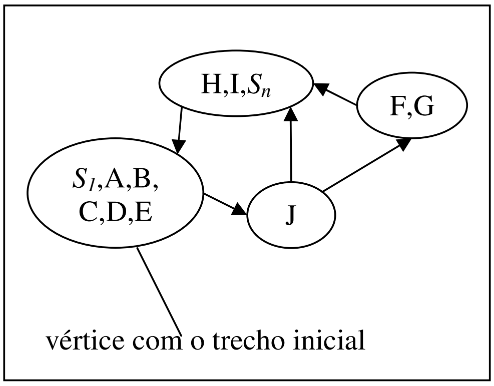


Figura 2 – O grafo como representação computacional da rota da figura 1.

3. Algoritmo para aprendizagem de rotas

O algoritmo para aprendizagem de rotas deve ser capaz de identificar os trechos que compõem a rota e as restrições temporais associadas a cada trecho ou a cada conjunto de trechos. Para detectar uma rota, o algoritmo possui duas fases: na primeira fase, o caminho que o veículo percorreu é detectado, onde cada trecho é adicionado ao caminho com seu intervalo de tempo; na segunda fase, a rota é atualizada com o caminho que foi detectado na fase anterior. No que se segue, consideraremos apenas rotas circulares.

A detecção de um caminho percorrido começa e termina nas seguintes condições. Quando o ônibus sai da garagem, antes que ele chegue à posição inicial de viagem, o ônibus permanece no estado “fora de serviço”. Assim que o ônibus chega a sua posição inicial de viagem, o estado do ônibus muda para o estado “em serviço”. Assim que o ônibus entra no estado “em serviço”, a detecção do caminho deste ônibus é iniciada. Ao chegar novamente à posição de início de viagem, a detecção do caminho termina e começa novamente. Ao final da jornada, o motorista deve indicar que o trabalho foi terminado naquele dia.

A composição de um caminho é feita através de tuplas de segmento e intervalo de tempo permitindo que a mesma instância de segmento possa fazer parte de um outro caminho, possivelmente associado a outro intervalo de tempo. A detecção do caminho é um algoritmo de atualização (figura Figura 3) que a cada sinal GPS (GPSData) coleta o trecho mais indicado em relação ao sinal GPS recebido (linha 3) e atualiza a composição do caminho (linha 5). A atualização da composição do caminho depende de o trecho coletado já pertencer ou não ao caminho que está sendo detectado. Se o trecho já pertencer ao caminho, apenas o tempo final do intervalo de tempo do trecho é atualizado no caminho (linha 10). Se o trecho ainda não pertencer ao caminho, o trecho é adicionado (linha 12) e seu intervalo de tempo é iniciado com o tempo inicial e final iguais ao tempo do sinal GPS.

```
1  updatePath(Path path, GpsData gpsData) {
2      Segment lastSegment = path.getLastSegment();
3      Segment segment = CollectSegment(gpsData, lastSegment);
4      Time time = gpsData.getTime();
5      path.update(segment, time);
6  }
7
8  updatePathSegment(Segment segment, Time time) {
9      if (this.contain(segment)) {
10         this.updateSegmentTimeInterval(segment, time);
11     } else {
12         this.addSegment(segment, time);
13     }
14 }
```

Figura 3 – Algoritmo de detecção do caminho percorrido por um ônibus.

Ao finalizar a detecção do caminho, o próximo passo é atualizar a rota. O algoritmo de atualização da rota está ilustrado na figura 4. No algoritmo, ao chegar o caminho detectado, se a rota ainda não possuir nenhum caminho, todos os trechos e seus respectivos intervalos temporais do caminho detectado são adicionados em um único vértice e este vértice é adicionado ao grafo que representa a rota (linha 6 até a linha 12). Se a rota já possuir algum caminho, então o algoritmo seleciona cada trecho e o respectivo intervalo de tempo do caminho detectado e verifica se o trecho já foi adicionado na rota (linha 18 a linha 37). Se o trecho ainda não foi adicionado, ele é adicionado com seu intervalo de tempo e cria-se uma aresta entre o último trecho verificado (se existir) e o que acabou de ser adicionado. Se o trecho já foi adicionado, o seu intervalo de tempo é atualizado na rota e verifica se já existe uma aresta entre o último trecho verificado (se existir) e o trecho que acabou de ter seu intervalo de tempo atualizado. Se a aresta não existir, então é adicionada a aresta com o último trecho verificado (se existir) e o último trecho que teve seu intervalo de tempo atualizado. Ao atualizar a rota quando ela já possui um caminho, é necessário expandir os vértices para facilitar o algoritmo. A expansão dos vértices separa cada trecho em um vértice separado mantendo a ordem entre os trechos. Ao finalizar a atualização a rota é colapsada, colocando os trechos sebsequentes sem bifurcação nem encontro dentro de um mesmo vértice.

4. Testes

Para testar o algoritmo para aprendizagem de rotas, foram utilizados dados artificiais criados por um gerador de coordenadas com *timestamp* (para cada coordenada, um *timestamp*). Dado um conjunto de identificadores de trechos de rua e o número de coordenadas por trecho de rua, o gerador retorna um conjunto de coordenadas com *timestamp*.

Para simular a imprecisão do GPS, o gerador de coordenadas com *timestamp* utiliza uma imprecisão opcional que é aplicada às coordenadas geradas. Como pode ser visto na figura Figura 5, as coordenadas são geradas de maneira imprecisa sendo posicionadas a uma certa distância dos trechos das ruas. A imprecisão pode ser negativa ou positiva, determinando de qual lado do trecho da rua a coordenada será posicionada.

```
updateRoute(Path path) {
    if (path.size() == 0) {
        return;
    }
    if (this.size() == 0) {
        PathObject firstObject = path.getObject(0);
        Vertex vertex = new Vertex();
        for (int i = 0; i < path.size(); i++) {
            PathObject pathObject = path.getObject(i);
            vertex.addPathObject(pathObject);
        }
        this.addVertex(vertex);
    } else {
        this.expand();
        Vertex lastVertex = null;
        for (int i = 0; i < path.size(); i++)
        {
            PathObject pathObject = path.getObject(i);
            Vertex vertex = this.getVertex(pathObject);
            if (vertex != null) {
                vertex.update(pathObject);
                if (lastVertex != null) {
                    if (!this.containEdge(lastVertex, vertex)) {
                        Edge edge = new Edge(lastVertex, vertex);
                        this.addEdge(edge);
                    }
                }
            } else {
                vertex = new Vertex();
                vertex.addPathObject(pathObject);
                this.addVertex(vertex);
                if (lastVertex != null) {
                    Edge edge = new Edge(lastVertex, vertex);
                    this.addEdge(edge);
                }
            }
            lastVertex = vertex;
        }
        this.collapse();
    }
}
```

Figura 4 – Algoritmo de atualização da rota.

A imprecisão do GPS pode causar alguns problemas aos sistemas que realizam processamento com esses dados. A figura 5 ilustra um desses problemas. Um dispositivo GPS em um ônibus que trafega a rua na parte superior da figura, da esquerda para a direita, gera sinais GPS com uma determinada imprecisão até que um ponto acaba ficando, incorretamente, mais próximo da rua na parte inferior da figura. Ao coletar os trechos de rua, armazenados no banco de dados, mais próximos da coordenada do sinal GPS em questão, o trecho de rua mais indicado para ser o escolhido seria o trecho da rua na parte inferior da figura por ser o trecho de rua mais próximo da coordenada atual. Mas, neste caso, o algoritmo de coleta dos trechos leva em consideração o último trecho de rua coletado e o primeiro critério para escolha do trecho é a continuação do trecho anterior. Assim, dos trechos coletados da base de dados mais próximos da posição do ônibus, se algum deles possuir conexão com o último trecho de rua, este será o trecho de rua escolhido; caso contrário, o trecho de rua mais próximo da posição do ônibus será o escolhido.

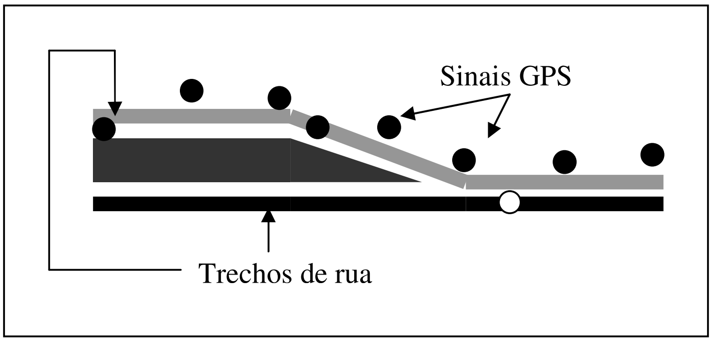


Figura 5 – Visualização da imprecisão das coordenadas geradas pelo gerador de coordenadas.

5. Conclusões

Este trabalho apresentou inicialmente um modelo para rotas. Em seguida, descreveu brevemente um algoritmo para aprendizagem automática de rotas de ônibus. Por fim, discutiu alguns aspectos práticos para detecção de trechos decorrentes dos erros inerentes ao uso de GPS. O trabalho encontra-se em estágio preliminar e precisa ser estendido para detecção de rotas bidirecionais e para acomodar variações sazonais de rotas mais complexas, dependentes da época do ano (carnaval, independência, etc) ou de eventos pontuais (acidentes, obras, etc), entre outros.

Referências

[1] Vaarandi, R., "SEC - a lightweight event correlation tool". IEEE Workshop on IP Operations and Management, pp. 111-115, 2002.

[2] Kruegel, C., Toth, T., Kerer, C., "Decentralized Event Correlation for Intrusion Detection". Proceedings ICISC 2001.

[3] Clemente, R. G., "Uma Arquitetura para Processamento de Eventos de Log em Tempo Real". Dissertação de Mestrado, Departamento de Informática da Pontifícia Universidade Católica do Rio de Janeiro, 2008.

[4] Symes, D. J., "Automatic Vehicle Monitoring: A Tool for Vehicle Fleet Operations". IEEE Transactions on Vehicular Technology, vol. VT-29, No.2, May 1980.

[5] Hornsby, K. S., "Modeling Moving Objects: Future Research Possibilities". Colloquium for Andrew U. Frank's 60th Birthday, 2008.

[6] Jakobson, G., Weissman, M., Brenner, L., Lafond, C., Matheus, C., "GRACE: Building Next Generation Event Correlation Services". Proceedings of the 7th Network Operations and Management Symposium, pp. 701-714, April 2000.

Técnica de Fusão de Imagens para Facilitar a Detecção de Áreas Canavieiras em Incompatibilidade Ambiental

Maikon R. de Novaes[1], **Bernardo T. F. Rudorff** [1], **Edison Crepani**[1]
Luciana Miura Sugawara[1]

[1]Divisão de Sensoriamento Remoto - DSR
Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12201-970 - São José dos Campos - SP, Brasil

{maikon,bernardo,crepani,lmiura}@dsr.inpe.br

Abstract: *The purpose of this paper is to develop a methodology based on the techniques of GIS and remote sensing as the fusion techniques of images to identify cane crops in APP in the São Paulo state. For such images were used pancromatic the HRC (High Resolution Camera) and multispectral CCD (Charge Coulpe Device) data. The intersection between APP and Sugarcane crop maps, combined with a careful visual interpretation with the aid of the fusion image, resulting in the areas of legal conflict. In the study area was found 9,840 hectares of APP in 996,100 hectares of sugarcane crop of which 60 hectares have found themselves in legal incompatibility environment. Fusion images technique allowed sugarcane crops identification as well as the detailed layout of the drainage network in the next scale 1:10.000.*

1. Introdução

As Áreas de Preservação Permanente são definidas por lei como áreas protegidas com a função ambiental de preservar os recursos naturais. Com isto os direitos de propriedade devem ser exercidos em conformidade com as limitações que a legislação em geral, e especialmente a Lei nº. 4.771 de 15 de setembro de 1965, estabelecem no Código Florestal.

Considerando a necessidade de regulamentar o Art. 2º da referida Lei foram elaboradas as resoluções CONAMA nº 302 (reservatórios artificiais) e nº 303 (cursos d'água e nascente). Embora estas resoluções estabeleçam os parâmetros, as definições e os limites das APP não se dispõem atualmente de mapas que adequadamente representem estes espaços territoriais protegidos tampouco o uso e ocupação no entorno dos reservatórios artificiais, dos cursos d'água e das nascentes.

As imagens de sensoriamento remoto têm ampliado bastante o horizonte de suas aplicações nos últimos anos devido à melhor qualidade e maior disponibilidade das imagens permitindo observar a superfície terrestre com maior detalhe e maior freqüência. A integração destas imagens com outras fontes de dados espaciais em Sistemas de Informação Geográfica (SIG) e Banco de Dados Geográficos (BDG) ampliaram significativamente o leque de possibilidades de análises espaciais de dados da superfície terrestre.

A fim de facilitar a extração de dados de imagens orbitais pelos usuários pode-se contar com diversas técnicas de processamento de digital de imagens, isto é, a análise e o tratamento de imagens por computador. Uma destas técnicas amplamente utilizadas é a fusão de imagens. Esta consiste em unir a informação de bandas multiespectrais de diferentes faixas de comprimento de onda com a melhor resolução espacial presente nas bandas pancromáticas dos sensores (Schowengerdt, 2007) a fim que os alvos terrestres sejam mais distinguíveis pelo olho humano, possibilitando extrair informações em escalas maiores de trabalho.

Diante de novos lançamentos de satélite, avanços na tecnologia de sensores orbitais e crescente disponibilidade de imagens de satélite de alta resolução espacial para uso civil surgem novas possibilidades de usar técnicas de processamento digital de imagens e geoprocessamento a fim de identificar áreas de preservação permanente dispostas na legislação ambiental, bem como o seu uso presente.

Neste sentido o presente trabalho tem por objetivo utilizar as mais recentes geotecnologias disponíveis para gerar um mapa das Áreas de Preservação Permanente de cursos d'água e identificar a incompatibilidade legal da ocupação destas APP com o cultivo de cana em uma área piloto situada no norte do Estado de São Paulo. Devido à indisponibilidade de uma base vetorial de hidrografia em escala próxima a 1:10.000, buscar-se-á, neste trabalho, extrair uma rede de drenagem em escala equivalente de detalhes.

2. Material Utilizado

Para este trabalho foi utilizada uma imagem adquirida pelo sensor CCD – órbita/ponto 158/123 e duas do sensor de alta resolução HRC – sub-órbita/ponto 158_C/128_1 e 158_C/128_2 correspondendo à data de 28 de junho de 2008.

Além destas foram adquiridas uma imagem Landsat-7/ETM+ Geocover (NASA, 2000)[1], georeferenciada, que servirá de base para o registro das imagens do sensor CCD (Charge Coulpe Device) e HRC (High Resolution Camera) com resoluções espaciais de 20 m e 2,7 m, respectivamente. Somando-se a estas foram utilizadas bases cartográficas vetoriais pré-existentes de municípios (IBGE, 2005); de drenagem escala 1:50.000 (IBGE, 2008) e uma imagem da missão SRTM - *Shuttle Radar Topography Misson* – (USGS, 2008) reamostrados os 90 m para tamanhos de pixels de 30 m por Valeriano (2004).

Um mapa temático com áreas agrícolas mapeadas de uso de cana-de-açúcar produzidas pelo projeto CANASAT[2] serviu para configurar, posteriormente, o mapa de incompatibilidade legal.

Foram utilizados aplicativos como o ESRI ArcGIS 9.2®, e seu *plugin* ArcHydro 1.2 (ESRI, 2007) para extração de rede de drenagem. Para outros processamentos de imagens, bem como armazenamento do banco de dados adotou-se o aplicativo SPRING 4.3.3 (Câmara et al., 1996).

3. Metodologia

3.1. **Pré-processamento**

---

[1] Disponível via http://zulu.ssc.nasa.gov/mrsid/. Acessado em julho de 2008.

[2] Disponível via http://www.dsr.inpe.br/canasat/. Acessado em julho de 2008.

Adquiridas as cenas citadas anteriormente (Landsat-7/ETM+ Geocover; CBERS-2B/CCD e CBERS-2B-HRC) foi definida uma área para o projeto correspondente à área das duas cenas do sensor de alta resolução – HRC – compreendida entre as longitudes W 49° 47' 56.60" - 49° 27' 50.95" e latitudes S 20° 23' 5.94" – 20° 4' 28.52". A imagem georreferenciada ETM+ Geocover serviu como referência para o registro sendo que as duas cenas HRC compõem um mosaico abrangendo partes dos municípios de Riolândia, Paulo de Freire, Pontes Gestal, Palestina, Américo de Campos, Cosmorama e Tanabi, todos inseridos no estado de São Paulo.

3.2. Fusão HRC-CCD

As técnicas de fusão permitem integrar a melhor resolução espacial da banda pancromática à melhor resolução espectral das demais bandas, produzindo uma imagem colorida que reúne ambas as características. As técnicas de fusão incluem: Intensidade-Matiz-Saturação (IHS), Principais Componentes, Transformação Wavelet, dentre outros (Schowengerdt, 2007; Gonzales, 2001). No entanto, o método de IHS é um dos mais utilizados devido a sua eficiência e facilidade de implementação (Tu et al., 2001) sendo aplicado, também, para extração de feições hidrológicas (Renzulo et al., 2008). Logo, optou-se por uma fusão por transformação IHS onde o resultado é apresentado na Figura 1 (a,b).

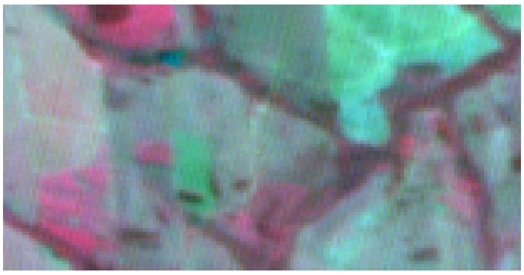


Figura 1a – Imagem CCD - R4B2G3

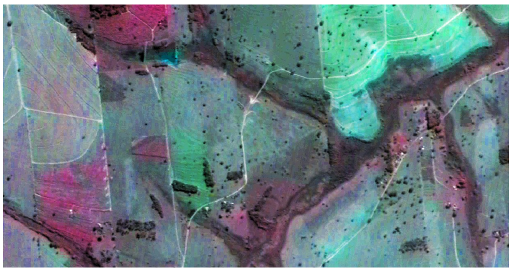


Figura 1b – Imagem Fusão (CCD-HRC) - R4B2G3

3.3. Criação da Rede de Drenagem e Mapa de APP

Primeiramente, foi importado um mapa contendo a rede de drenagem escala 1:50.000 a fim de gerar, em seguida, um *buffer* contendo as áreas de distâncias. Todavia, devido à resolução espacial fina proporcionado pela fusão das imagens, o mapa pré-existente tornou-se incompatível com a escala de trabalho.

Diversos trabalhos (Valeriano, 2004; Wright et al., 2006) mostram que Modelos Digitais de Elevação (MDE) são de grande valia quando se utiliza Sistemas de Informações Geográficas (SIG) em Cartografia Regulamentar, auxiliando no ordenamento do território, planejamento urbano e regional, zoneamento, quantificação de grandezas ligadas às características físicas da bacia e identificação da rede de drenagem e divisores de água.

O radar de abertura sintética (SAR) por interferometria mostra-se muito eficaz para medir a topografia digital. A Missão Topográfica por Radar interferométrico, ou Shuttle Radar Topography Mission (SRTM), gerou um modelo digital de elevação (MDE) de resolução de 90 m, interpolados para 30 m (Valeriano, 2004).

Através do algoritmo ArcHydro 1.2 do aplicativo ArcGIS 9.2 fez-se três extrações automáticas a partir do MDE do SRTM usando os seguintes parâmetros de segmentação: 200, 1000 e 1500 pixels. A fim de buscar o melhor resultado que fosse compatível com a escala próxima de 1:10.000 foi utilizado como auxílio a rede gerada a partir de 1000 pixels pois esta não superestimou as áreas de drenagem e mostrou um melhor resultado visual quando inserida sobre a imagem fusionada.

A imagem fusionada aliada aos dados de extração automática da rede de drenagem foram utilizados como base para o traçado manual da rede de drenagem e posterior obtenção do mapa de distâncias (*buffers)* referentes às APPs seguindo a definição da legislação ambiental (Figura 2).

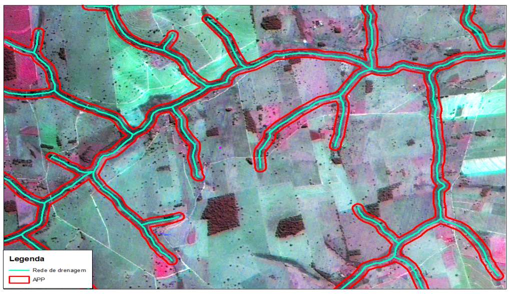


Figura 2 – Imagem fusionada com o traçado manual da rede de drenagem e APPs

Um mapa de distância é um tipo de análise de proximidade (medida de distância entre objetos), comumente medida em unidade de comprimento, que apresenta zonas com larguras especificadas (distâncias) em torno de um ou mais elementos de um mapa temático vetorial a partir da definição de fatias associadas a classes temáticas previamente definidas no banco de dados. Um mapa de distância ou “buffer” é definido como sendo uma área de extensão regular, que é desenhada ao redor de um ou mais elementos espacialmente definidos (pontos, linhas ou polígonos) (INPE, 2000).

As APPs são determinadas em legislação a partir da definição de fatias correspondentes à área situada em faixa marginal do curso d’água, medida a partir do nível mais alto (nível alcançado por ocasião da cheia sazonal do curso de água perene ou intermitente), em projeção horizontal, com largura mínima de: (i) 30 metros para curso de água com menos de 10 metros de largura; (ii) 50 metros para curso de água com 10 a 50 metros de largura; (iii) 100 metros para curso de água com 50 a 200 metros de largura; (iv) 200 metros para curso de água com 200 a 600 metros de largura e (v) 500 metros para curso de água com mais de 600 metros de largura.

Os lagos e lagoas naturais foram também vetorizados a partir da imagem fusionada com auxílio da rede de drenagem extraída da imagem SRTM e foram acrescidas ao seu redor uma faixa de 50 metros quando localizados em área rural e apresentarem até vinte hectares de superfície e faixa de 100 metros para os demais localizados em área rural.

3.4. Áreas de Incompatibilidade Legal para a Cana.

O mapa das APPs foi sobreposto a um mapa temático da cana-de-açúcar obtido por meio de imagens multitemporais do sensor TM do satélite Landsat-5. A intersecção dos dois mapas, aliado a uma cuidadosa interpretação visual com o auxílio da imagem fusionada, resulta nas áreas de incompatibilidade legal.

As áreas de incompatibilidade legal para drenagem correspondem às Áreas de Preservação Permanente que foram utilizadas para alguma atividade diferente daquela prevista em lei.

4. Resultados e Considerações

Verificou-se que a área de estudo tem 9.840 ha de APP e 996.100 ha cultivados com cana dos quais 60 ha estão inseridas em APP, ou seja, em incompatibilidade legal. O método de fusão de imagens favoreceu a identificação de lavouras de cana-de-açúcar bem como o traçado detalhado da rede de drenagem em escala próxima a 1:10.000 (Figura 2).

No entanto, deve-se atentar que os mapas produzidos estão em escalas diferentes de trabalho. O mapa realizado no projeto CANASAT trabalha em escalas de resolução compatíveis com o produto do Landsat/TM, de 30m. Por outro lado, o mapa aqui produzido possui uma escala mais refinada, de 2,7 m o que possibilita o trabalho em escalas até 1:5.000.

Devido à pouca disponibilidade de mapas em escala compatível realizada neste trabalho, a validação do mesmo deve ser efetuada por meio de um trabalho de campo onde devem ser sorteados pontos amostrais, aleatoriamente, na área das APP geradas a fim de confirmar o sucesso desta metodologia. Com a disponibilidade de imagens de alta resolução combinadas com técnicas de processamento de imagem como a fusão pode-se realizar trabalhos de mapeamento, zoneamento e monitoramento ambiental em escalas superiores a 1:10.000 de maneira eficiente e dinâmica.

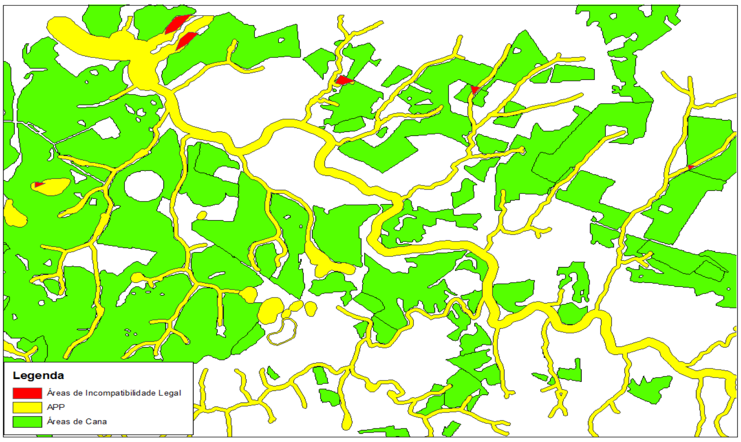


Figura 3 – Áreas de lavoura de cana-de-açúcar em incompatibilidade legal.

As imagens de alta resolução também contribuem para diminuir possíveis dúvidas durante a prática de classificação e interpretação de imagens no processo de mapeamento agrícola das lavouras de cana-de-açúcar. Estas podem, desta forma, auxiliar no mapeamento destas áreas contribuindo para estimativas mais precisas de safras.

Apesar da fusão ter apresentado um bom resultado, recomenda-se para este tipo de técnica um melhor registro das cenas. No caso de imagens de alta resolução é necessário adquirir pontos em campo, através de um GPS de precisão, inferior a um metro. Com isso, pode-se ter mais propriedade na confirmação de áreas de incompatibilidade legal de uso.

Para determinados fins, mapeamentos ou classificações automáticas ainda não respondem de maneira realística o que torna tal trabalho manual uma operação penosa e de grande custo para o operador. Assim, esta pesquisa abre campo para novas possibilidades de trabalho no mapeamento de áreas de preservação permanente a fim de que estas não se tornem alvos de práticas incompatíveis de uso e perante a lei.

5. Referências

Câmara, G.; Souza, R. C. M.; Freitas, U. M.; Garrido, J. C. P. (1996) "SPRING: Integrating Remote Sensing and GIS with Object-Oriented Data Modelling". Computers and Graphics, v. 15, n. 6, p. 13-22.

Constituição Federal - Lei nº. 4.771 de 15 de setembro de 1965. (2004) *in:* "Coletânea de Legislação de Direito Ambiental". Organizador Odete Medauar. 3.ed. rev., atual. e ampl. São Paulo: RT.

ESRI (2007). "Arc Hydro Tools version 1.2 for ArcGIS 9.2 Tutorial". California. 110 p.

Gonzalez, R. C.; Woods, R. E. (2000) "Processamento de Imagens Digitais". São Paulo: Ed. Edgard Blücher, 509 p.

Instituto Brasileiro de Geografia e Estatística – IBGE. (2005)"Malha Municipal do Brasil: Situação 2005". Rio de Janeiro: IBGE.

INPE – Instituto Nacional de Pesquisas Espaciais. (2000) "Fundamentos de Geoprocessamento". Tutorial – SPRING.

Renzullo, L. J., Barrett, D. J., Marks, A. S., Hill, M. J., Guerschman, J. P., Mu, Q., Running, S. W. (2008). Multi-sensor model-data fusion for estimation of hydrologic and energy flux parameters. Remote Sensing of Environment. v. 112, p.1306–1319.

Schowengerdt, R. A. (2007) "Remote sensing models and methods for image processing". San Diego: Academic Press, 515 p.

Tu, T.; Su, S.; Shyu, H.; Huang, P. S. (2001) "Efficient intensity-hue-saturation-based image fusion with saturation compensation". Optical Engineering, v. 40, n. 5.

USGS. United States Geological Survey. (2008) "Shuttle Radar Topography Mission (SRTM)".<http://edc.usgs.gov/products/elevation/srtm>. Acesso em agosto de 2008.

Valeriano, M. M. (2004) "Modelo digital de elevação com dados SRTM disponíveis para a América do Sul". São José dos Campos, SP INPE: Coordenação de Ensino, Documentação e Programas Especiais (INPE-10550-RPQ/756). 72p.

Wright, R.; Garbeil, H; Bologa, S. M.; Mouginis-Mark, P. J. (2006) "An assesment of shuttle radar topography mission digital elevation data for studies of volcanomorphology". Remote Sensing of Environment. v. 105. p 41-53.

Spatial Patterns of Residential Segregation:

A Generative Model

Flavia F. Feitosa[1], Joshua Reyes[2], Walter E. Zesk[3]

[1]Center for Development Research (ZEF) - University of Bonn – Bonn, Germany

[2]Department of Computer Science – University of Massachusetts Boston - USA

[3]Rhode Island School of Design and Architecture - USA

flavia@uni-bonn.de, jreyes@cs.umb.edu, wzesk@wesleyan.edu

Abstract. *Residential segregation emerges from the interaction of many individuals and displays makedly different global patterns depending on specific socioeconomic contexts. This paper presents a generative model of socioeconomic segregation that reproduces regular macro-patterns of the phenomenon through the specification of a minimal amount of parameters driving the agent's behaviour. The purpose of these experiments is to provide insights about the emergence of certain types of segregation patterns that have been identified in many modern cities and measure the relation between these distinct outputs and the degree of segregation they produce.*

1. Introduction

Residential segregation is a measure of social clumping in an urban environment. It has different meanings depending on the specific form and structure of the city, and its categories include income, class, race, and ethnic segregation. The effects of segregation in a city are mainly negative. In particular, socioeconomic segregation limits access of disenfranchised population groups to infrastructure and job opportunities, while reinforcing racial and social prejudices [Feitosa, *et al.* 2007; Sabatini, *et al.* 2001].

Residential segregation exhibits many of the characteristics hallmarks of complex adaptive systems, particularly emergence and non-linearity. Segregation is perceived as a large-scale urban phenomenon, but emerges from the interaction between individuals at a local level [Schelling 1978]. Positive feedbacks introduce a non-linearity into the system. As a result, small differences in socioeconomic and physical context and local behaviour can generate large, unexpected and sometimes counter-intuitive outcomes that cannot be understood as the simple sum of the constituent parts [Gilbert 2008]. Segregation occurs in most large modern cities, including the developed and the developing word. An examination of the patterns of segregation in various cities reveals the impact of specific socioeconomic environments. For example, most large cities in the United States have poor and non-white ghettos as a persistent feature of

This work was partially supported by the Santa Fe Institute whose research and education programs are supported by core funding from the National Science Foundation and by gifts and grants from individuals, corporations, other foundations, and members of the Institute's Business Network for Complex Systems Research.

their central areas, whereas the wealthy and mostly white population prefers to live in suburbs [Massey and Denton 1993]. In Latin America, wealthy families concentrate in areas that expand from the historical centre into a single geographical direction, while the poorest population mostly settle in the roughly equipped peripheries [Sabatini, *et al.* 2001; Villaça 2001]. These two outcomes are generally recognized, but the structural aspects that cause these distinct patterns remain unclear. What part do individual needs and preferences play in the mechanics of segregation? How do socio-demographic factors, such as income inequality, impact the emergence of different spatial outcomes?

Generative models have been used to understand issues related to structural characteristics of complex social systems [Epstein 2007]. Most often, an agent-based computational model attempts to reproduce certain global regularities from minimally defined local interactions among individuals within the system. Such a model makes it possible to conduct experiments on a simplified population of agents in a spatial environment. Agents interact locally according to small set of rules, thereby generating global regularity from the bottom up. Simply, the motto of generative social approach is: 'If you didn't grow it, you didn't explain it' [Epstein 2007].

This paper presents a generative model of socioeconomic segregation that reproduces patterns much like those observed in real cities. The rules defining the experiment are few and simple. The purpose of this experiment is to create a stylized, abstract model that is able to provide insight into the mechanics of emergence and the variety of segregation patterns identified in modern cities. In addition, this work understands each pattern in terms of the degree of segregation produced. Our simulation enables the quantitative comparison of different patterns of segregation on defined segregation measures.

2. The Model Specification

Our model relies on the premise that socioeconomic segregation is the *outcome of a contest for the most convenient locations in the city.* In this model, only two attributes define how convenient a location is: the *quality* and *proximity* to the central business district (CBD). The environment consists of a two-dimensional 45 X 45 grid. Each cell within the grid corresponds to a dwelling unit. At the beginning of the run, each cell receives an equal quality rating. The initial price distribution falls of as the inverse of distance to the CBD.

Since the aim of our model is to 'grow' a city, the grid is initially populated with a minimum amount of agents. Each agent corresponds to a household, which is characterized by its location and income. Incomes are drawn according to a Pareto distribution, the skewness of which can be controlled by its exponent, known as Pareto index. The Pareto index regulates the level of income disparity in the city and can be chosen by the user: higher values lead to an unequal city, and vice-versa.

The city is gradually populated by households according to a growth rate. While looking for their residential location, the new inhabitants of the city attempt to maximize the utility *U* derived by a location over some subset of the presently unoccupied cells according to the Cobb-Douglas function:

$$U = \left(\frac{1}{D}\right)^{1-\alpha} Q^{1+\alpha},$$

where Q is the quality index of the cell, D is the distance to the CBD, and $\alpha \in [-1,1]$ is the quality-priority index. The quality-priority index introduces a cultural bias to the model. Households will prioritize locations with high quality if α is near 1, and prioritize cell close to the city's center for values of α close to -1.

The individuals in this model exhibit bounded rational behavior. Because search is costly, every household evaluates a limited number of the total, possible locations to move to. Unlike other models of segregation, the households choose their new location without giving consideration their neighbours or their income explicitly. At each update, however, individuals compare utilities among others with similar level of income. Households with utilities that are much lower than those with similar income grow dissatisfied and may repeat the search process and move to another location.

As individuals choose their residential location, they change the environment. The model assumes that individuals with high income are able to advocate for more public goods in their neighbourhood, and therefore slightly increase property value (price) and introduce more services and facilities, such as improved school systems, police, or public parks (quality). On the other hand, individuals with low income lack similar resources, and so property value and public goods in those neighbourhoods decrease slightly.

The spatial arrangement of the population is constantly monitored through segregation indices. We selected segregation indices that are able to capture the segregation dimensions defined by Reardon and O'Sullivan [2004]: *spatial evenness/clustering* and *spatial exposure/isolation*. To measure each of these dimensions, we compute spatial indices of segregation proposed by Feitosa et al. [2007].

Spatial evenness/clustering refers to the balance of the distribution of population groups, and it is measured by the *generalized neighbourhood sorting index* $GNSI$. The *GNSI* evaluates how much of the income variance between different neighbourhoods contributes to the total income variance of the city. Generally, higher variance between neighbourhoods signifies higher levels of segregation along the dimension evenness/clustering. Spatial exposure/isolation refers to the likelihood for members from different groups to live side-by-side, and it is captured by the *spatial exposure index* $\breve{P}_{(m,n)}$. The $\breve{P}_{(m,n)}$ measures the average proportion of group *n* in the neighbourhoods of each member of group *m*. Since both these indices are global quantities, we also compute local indices of spatial isolation ($\breve{q}_m$) that can be displayed as maps and facilitate the visual interpretation of the simulations [Feitosa, *et al.* 2007].

3. Growing Spatial Segregation Patterns

We ran our simulation several times for varying preferences and economic background. During these runs the quality-priority ranged over the values α = -1, -0.5, 0, and +0.5 for each scenario captured by the Pareto inequality indices β = 0.5, 1, and 1.5 – cities with low, moderate, and extreme disparities in wealth. Each simulation ran until its population exceeded one thousand households.

Figure 1 presents the distribution of households for two runs in our experiment. The rightmost column depicts the physical distribution of households, colored by income, across the model's landscape. In this picture, squares with different tones of red, orange, and blue correspond to households with low, moderate, and high incomes respectively. In the middle of the rightmost columns, we have plotted the spatial isolation of the poor ($\breve{q}_{POOR}$) and of the wealthy ($\breve{q}_{WEALTHY}$). Dark cells signify areas where these groups are exceptionally isolated.

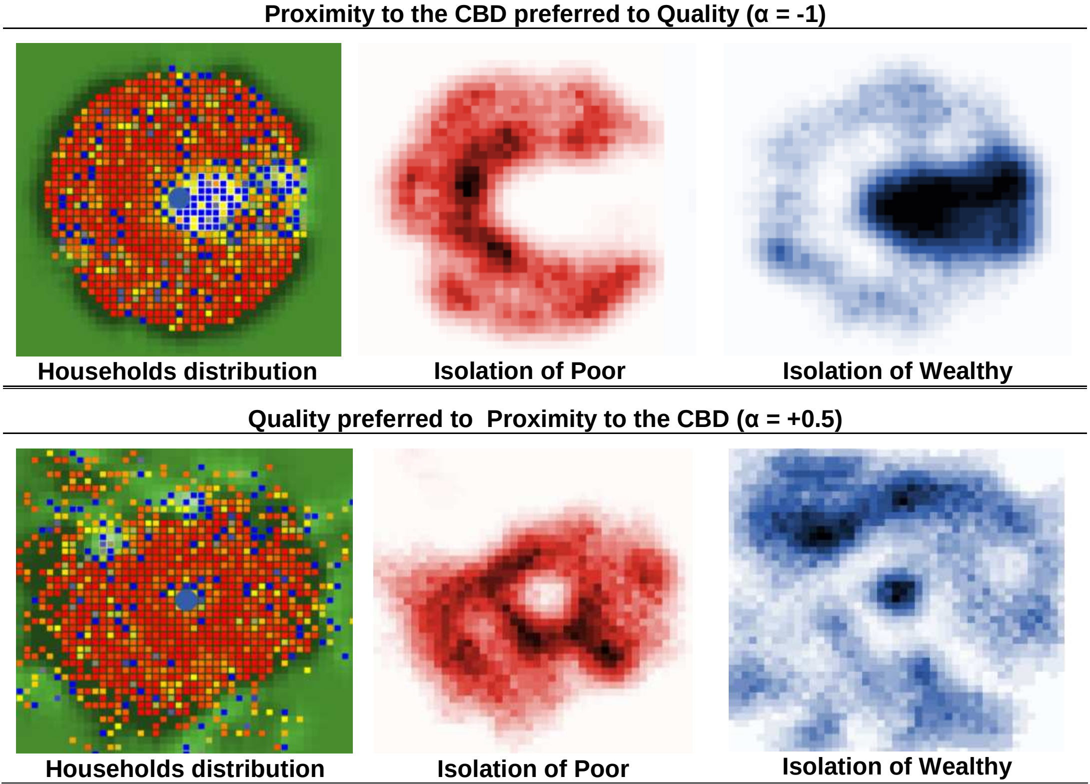


Figure 1. Simulation outcomes based on different values for the quality-index *α*.

We find that negative values of the quality priority index *α*, that is, the proximity to the CBD is preferred to quality, generate a very compact city. This output resembles a classical pattern of segregation known as Hoyt's sector model of segregation. According to Hoyt [1939], if a district is set up for high income residence, any new development in that district will expand from the outer edge and, therefore, the sector shape emerges. This sector pattern has been commonly observed in Latin American cities, where wealthy families concentrate in areas that expand from the center in a single direction. The top row in Figure 1 illustrates the case in this point.

On the other hand, positive values of *α*, that is, the quality is preferred to proximity to the CBD, generate a disperse city with wealthy suburbs. The bottom row in Figure 1 gives an example of this kind of city, which resembles a classical pattern of segregation known as Burgess's concentric model. For Burgess [1924], a city grows outward from a central point in a series of rings and there is a correlation between the distance from this central point and the wealth of residential areas. In North American cities, where lower transportation costs decrease the importance of proximity to the CBD, we commonly experience this pattern of urban development.

Figure 2 presents plots of segregation indices measured for simulations conducted with four different values for the quality-index α. In order to obtain more robust results, we repeated these measurements using five different random-seeds and three different degrees of inequality (presented in different colors). The first graphic shows how the index GNSI, which measures the segregation dimension evenness/clustering, presents a positive correlation with the quality-priority index. The second graphic shows how the exposure of wealthy families to poor families decreases with the increase in the quality-priority index. Both results indicate that as soon as households attribute less value to proximity and more value to quality, the segregation level increased in both dimensions. Therefore, the 'disperse city' dominated by wealthy suburbs seem to be the segregation pattern that produces the highest levels of segregation. Regarding inequality, the graphics of Figure 2 reveal an unexpected result: while higher income inequality increases the segregation in the dimension evenness/clustering (GNSI), it increases the exposure amongst different income groups and therefore decreases the segregation in the dimension exposure/isolation.

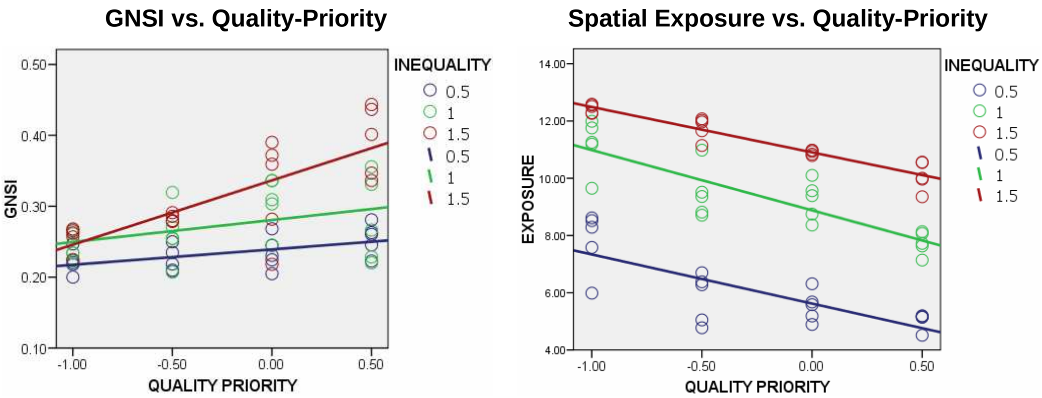


Figure 2. Segregation indices computed for simulation outcomes with different values for the quality-index *α*.

4. Concluding Remarks

This paper presents an agent-based model of segregation that generates patterns of segregation much like those identified in real cities. These patterns emerge despite the simplicity of the model. It is an example of how complexity can emerge from just a few parameters and suggests that these parameters are crucial to the underlying dynamics of segregation.

The model demonstrates how the variation of a single parameter, such as quality-priority was able to generate classical patterns of segregation. This is an indication of how cultural and economic environment influences segregation patterns. For instance, low transportation costs in North American cities minimize the importance of proximity to the center of the city and promoted the emergence of disperse cities with wealthy suburbs. It will be interesting to see the effect of the current energy crisis on segregation in these cities.

This work includes a quantitative comparison of segregation patterns through the application of global segregation indices. This is a particular advantage provided by generative models, since indices computed for real cities cannot be compared due to their susceptibility to the modifiable areal unit problem (MAUP). Our findings indicate that disperse cities produce higher degree of segregation in both dimensions, while unequal cities present higher degree of segregation in the dimension evenness/clustering, but not in the dimension exposure/isolation.

Segregation remains a remarkably common, persistent and difficult problem for cities of all types. Like many social phenomena, it grows out of simple interactions between individuals, but manifests as a complex resilient phenomenon at the scale of the city. Agent-based models provide a unique laboratory for experimentation and examination of this phenomenon, stripped down to its essential mechanics. We view this work as a simple kind of null model upon which urban planners and policy makers may introduce and isolate the effects of legislation, e.g., the social-mixed housing policies that have been recently instated in some North American cities.

5. References

Burgess, E. W. (1924), The growth of the city: an introduction to a research project, *Publications of the American Sociological Review*, *18*, 85-97.

Epstein, J. M. (2007), *Generative Social Science: Studies in Agent-Based Computational Modeling.*, Princeton University Press.

Feitosa, F. F., et al. (2007), Global and local indicators of urban segregation, *International Journal of Geographical Information Science*, *21*, 299-323.

Gilbert, N. (2008), *Agent-Based Models (Quantitative Applications in the Social Sciences)*, Sage Publications.

Hoyt, H. (1939), *The Structure and Growth of Residential Neighborhoods in American Cities*, Federal Housing Administration, Washington DC.

Massey, D. S., and N. A. Denton (1993), *American apartheid: segregation and the making of the underclass*, 291 pp., Harvard University Press, Cambridge.

Reardon, S., and D. O'Sullivan (2004), Measures of Spatial Segregation, *Sociological Methodology*, *34*, 121-162.

Sabatini, F., et al. (2001), Segregación residencial en las principales ciudades chilenas: tendencias de las tres últimas décadas y posibles cursos de acción., *EURE (Santiago)*, *27*, 21-42.

Schelling, T. C. (1978), *Micromotives And Macrobehavior*, W. W. Norton, London & New York.

Villaça, F. (2001), Segregation in the Brazilian Metropolis, paper presented at International Seminar on Segregation in the City, Lincoln Institute of Land Policy, Cambridge.

RDengue um ambiente para o monitoramento de ovos do mosquito *Aedes aegypti*

Wagner Hugo Bonat[1], **Henrique Silva Dallazuana**[1], **Paulo Justiniano Ribeiro Junior** [1], **Antônio Miguel Vieira Monteiro** [2]

[1]Laboratório de Estatística e Geoinformação – Universidade Federal do Paraná (LEG-UFPR)
CEP 81531-990 – Curitiba – PR – Brazil

[2]Instituto Nacional de Pesquisa Espacial (INPE)
São José dos Campos, SP – Brasil

{wagner,henrique,paulojus}@leg.ufpr.br, miguel@dpi.inpe.br

Abstract. *This work describes the development of an integrated environment providing support for monitoring amounts of eggs of the Aedes aegypti, the main vector of the dengue disease. The proposed environment combines computational resources and tools for spatial and/or temporal statistical analysis. The system is used on eggs count data obtained from the an experiment conducted within the context of the SAUDAVEL project. The environment integrates the statistical software R and the GIS tools Terralib and TerraView, using the package aRT and includes the development of RDengue, an R package implemented the proposed methodology and tailored for continuous an automatic surveillance systems.*

Resumo. *Este trabalho descreve o desenvolvimento de um ambiente integrado de apoio ao monitoramento de ovos de Aedes aegypti, o principal vetor da dengue. O ambiente proposto combina recursos computacionais e ferramentas para análise estatística espacial e/ou temporal. O sistema é utilizado em dados de contagens de ovos obtidos em um experimento conduzido dentro do contexto do projeto SAUDAVEL. O ambiente integra o programa estatístico R e as ferramentas de SIG Terralib e TerraView através do pacote aRT e inclui o desenvolvimento de um pacote do R, RDengue, que implementa a metodologia proposta, direcionada para o monitoramento contínuo e automático em um sistema de vigilância entomológica.*

1. Introdução

A degradação do meio-ambiente e os problemas sócio-culturais afetam o cenário epidemiológico brasileiro, levando-o a ser destaque na mídia nacional e internacional, decorrente de epidemias de dengue, leptospirose, a recorrência da tuberculose, entre outras. Diante disso, constatou-se a importância de utilizar sistemas de vigilância, capazes de detectar precocemente situações que caracterizam surtos epidêmicos, modelar e identificar fatores de risco e proteção nas situações endêmicas e epidêmicas. A detecção precoce de surtos por doenças transmissíveis como a dengue é importante para ativar ações de investigação e controle por partes das agências de saúde pública, o que reforça a necessidade de sistemas de vigilância. Nesta perspectiva, foi elaborado o ”Projeto SAUDAVEL”[Monteiro et al. 2006].

Segundo [Tauil 2002] a dengue é hoje a principal doença re-emergente no mundo. Na ausência de uma vacina preventiva eficaz, de tratamento etiológico e quimioprofilaxia efetivos, o único elo vulnerável para reduzir a sua transmissão é o mosquito *Aedes aegypti*, seu principal vetor.

Neste contexto, o objetivo deste trabalho é descrever o desenvolvimento de um ambiente R [R Development Core Team 2007] capaz de efetuar análises exploratórias, de forma automática e semi-automática. Tais análises, constituem um sistema de vigilância entomológica, dedicado ao monitoramento da atividade reprodutiva do mosquito *Aedes aegypti*. Através de contagens de ovos coletados com ovitrampas, em um experimento de campo conduzido na cidade de Recife/PE. Tal ambiente, acopla-se a um sistema de coleta de dados regular e continuada, e provê instrumentos para geração e disponibilização de resultados de análises dos dados em ambientes de fácil acesso para os profissionais envolvidos com o sistema.

Este sistema foi desenvolvido para os dados provenientes do experimento de coleta de ovos do mosquito *Aedes aegypti*, desenvolvido no escopo do projeto SAUDAVEL na cidade de Recife/PE. Aborda-se duas resoluções espaciais na análise: bairros e armadilhas.

Com o conjunto de rotinas propostas, é possível desenvolver análises espaciais e temporais, capazes de determinar limites de controle, verificar relações espaciais, identificar sazonalidades e tendências, estimar a ocorrência do fenômeno em áreas não observadas, definir áreas prioritárias para intervenção e também projetar possíveis cenários futuros, com o intuito de fornecer informações para subsidiar a definição de ações de controle do vetor da dengue.

2. Material e Métodos

Nesta seção serão descritos a área de estudo e os aspectos computacionais envolvidos na análise.

2.1. Área de estudo

O experimento de Recife/PE consiste de 564 armadilhas, distribuídas entre 6 dos 94 bairros da cidade e instaladas de modo a cobrir toda a superfície do bairro. As armadilhas começaram a ser monitoradas a partir de 03/2004, até a data de 05/2007 foram realizadas 19.068 coletas, nas quais foram contados 14.829.557 ovos do mosquito. A cada sete dias era feita a contagem em aproximadamente um quarto das armadilhas, assim em um ciclo de vinte e oito dias todas as armadilhas são monitoradas.

Neste experimento, cada armadilha contém uma lâmina na qual a fêmea do mosquito coloca os ovos, essas lâminas são recolhidas e a contagem é feita em laboratório especializado. Os dados são inseridos em um banco de dados geográfico através de uma interface *Web*, desenvolvida pelo INPE - Instituto Nacional de Pesquisas Espaciais, esta interface foi projetada para evitar formas complexas de entrada de dados. A descrição detalhada do experimento dentro da proposta da rede SAUDAVEL é apresentada em [Regis et al. 2008].

2.2. Aspectos computacionais: RDengue

Uma parte importante em projetos desta magnitude é a forma de armazenamento, tratamento e visualização dos dados. Um experimento como este gera uma grande quantidade de dados que não são facilmente manipuláveis, requerendo para isto ferramentas específicas para validação e análise. O banco de dados do Recife SAUDAVEL, está implementado segundo o modelo Terralib [Câmara et al. 2000], tecnologia de código aberto e MySql Database Server como um repositório e sistema gerenciador de dados espaço-temporais baseado no modelo espaço-temporal da Terralib [Silveira et al. 2004].

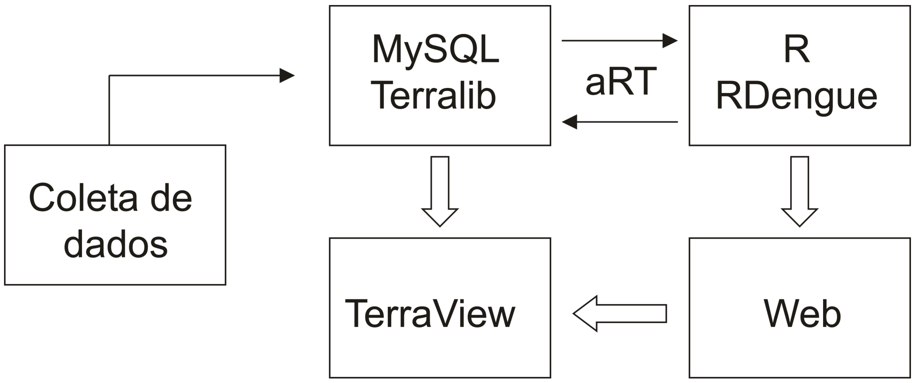


Figura 1. Formato geral de análise, as setas vazadas indicam visualização opcional, a setas sólidas indicam trocas persistentes.

A Figura 1 ilustra o formato geral de análise, iniciando pela obtenção dos dados e armazenagem em um banco de dados geográficos. Isto feito, pode-se conectar o banco a partir do ambiente R, através do pacote aRT [Andrade Neto et al. 2005]. Todas as análises estatísticas são realizadas em um ambiente especializado em implementações de métodos estatísticos, no caso o R, através das rotinas especializadas implementadas no RDengue. Após as análises serem concluídas, tem-se através do RDengue e do aRT a opção de retornar resultados das análises para o banco geográfico. Opções de disponibilização dos resultados incluem tabelas e camadas de informação adicionais no banco, sendo ainda possível, definir no ambiente R, opções de visualização a partir do acesso via uma ferramenta de SIG, como é o caso do TerraView, ou então gerar uma página *Web* para a visualização pública dos resultados. Os resultados são associados às unidades espaciais do experimento, bairros ou armadilhas. Tal integração de ferramentas torna possível disponibilizar de forma regular e automática os resultados e relatórios das análises, de forma personalizada ao usuário da informação, tornando transparentes, se desejado o ambiente computacional e estatístico subjacente. Ressalta-se, que sem essa integração de ferramentas seria difícil e custoso, proceder análises como as que serão posteriormente apresentadas. Porém, a ferramenta RDengue provê opções de análise em ambiente R com opções de disponibilização de resultados para outras ferramentas de SIG como o TerraView. As rotinas de análises implementadas no RDengue são exclusivas do ambiente, usando rotinas residentes do R aplicadas e/ou adaptadas para o caso do experimento de Recife/PE.

3. Resultados e discussões

O monitoramento da quantidade de ovos de *Aedes aegypti* visa a detecção precoce de surtos de dengue. Para isto, é necessário que métodos estatísticos monitorem momentos de anormalidades no crescimento da população do mosquito, distinguindo entre número de ovos capturados dentro do esperado, e uma situação em que o número de ovos sugere uma

ação imediata, com vistas a evitar ou reduzir o impacto de um possível surto da doença. Para atender a estes propósitos, o RDengue utiliza ferramentas estatísticas e de SIG, com o intuito de desenvolver análises espaciais e temporais, de forma automática e/ou semi-automática, tendo em vista a estrutura de coleta regular de dados. As possibilidades de uso de tais rotinas são amplas, sendo que estas são flexíveis e podem ser analisadas de diversas formas, além disso são de fácil interpretação por não especialistas.
A seguir, serão descritas como exemplo algumas opções de análises contidas no ambiente desenvolvido. A primeira opção corresponde a uma análise descritiva por armadilha, através de gráficos de controle e ajuste de modelos para séries temporais da classe *arima*[Hyndman and Khandakar 2008]. A segunda consiste em uma abordagem de análise de superfície. O pacote implementa ainda outras estratégias mais avançadas de modelagem que se incorporam ao ambiente, opta-se por não detalhar tais métodos aqui. O RDengue pode ser expandido acoplando metodologias que vêem sendo estudadas e avaliadas.

3.1. Análises por armadilhas

A Figura 2 consiste na captura de uma tela de análise ilustrando a explicita integração entre o R e o TerraView através da Terralib e do SGBD. Após gráficos para cada uma das 564 armadilhas serem produzidos pelas rotinas do RDengue, os resultados são acoplados ao banco de dados do SAUDAVEL através de uma tabela de *mídias*. Desta forma, uma ferramenta específica de SIG, no caso o TerraView, pode acessar os resultados da análise estatística produzida no R, com todas as facilidades e potencialidades que este ambiente tem para análise de dados. E a análise ganha o potencial de visualização própria de uma uma ferramenta de SIG como o Terraview. Além disso, as equipes de campo tem o acesso a tais análises muito facilitado, através do TerraView, o que não seria possível apenas através do R, pois demandaria maior familiaridade com a ferramenta estatística. O gráfico apresentado na Figura 2 utiliza as idéias das cartas de *Shewhart* [Montgomery 2000], para construir um gráfico de controle para a contagem de ovos do mosquito *Aedes aegypti*, possibilitando a comparação da situação de cada armadilha em relação as outras, ao bairro e também a cidade. O gráfico estabele limites de controle baseados em medidas de dispersão, para o nível local (bairro) e também global (município), apresenta as médias do município, do bairro, a média mensal e também a média das armadilhas vizinhas, a fim de verificar algum tipo de efeito espacial. Neste gráfico as coletas são agrupadas por mês. É possível também através destes gráficos identificar tendências e sazonalidades. Este tipo de gráfico é uma ferramenta simples, que permite obter informações comparativas sobre a evolução do número de ovos em cada armadilha. O RDengue possui ferramentas capazes de desenvolver este tipo de gráfico para todas as armadilhas, de forma automática e facilmente atualizável conforme a disponibilidade de dados de novas coletas, formando assim o sistema de vigilância e monitoramento desejado.
Um outro tipo de análise, considera as observações de cada armadilha como uma série temporal, ajustando um modelo da classe *arima*. A rotina do RDengue procura o modelo que melhor se ajusta as observações segundo o critério de *Akaike* [Akaike 1974]. São apresentados dois gráficos clássicos na identificação da ordem do modelo: o de autocorrelação e autocorrelação parcial. Esta análise identifica sazonalidade e tendências, permitindo, através do modelo ajustado, projetar um cenário futuro para a armadilha, tal como, predizer o número esperado de ovos para um ou dois meses a frente. Exemplos de tais análises, podem ser vistas no complemento online deste artigo.

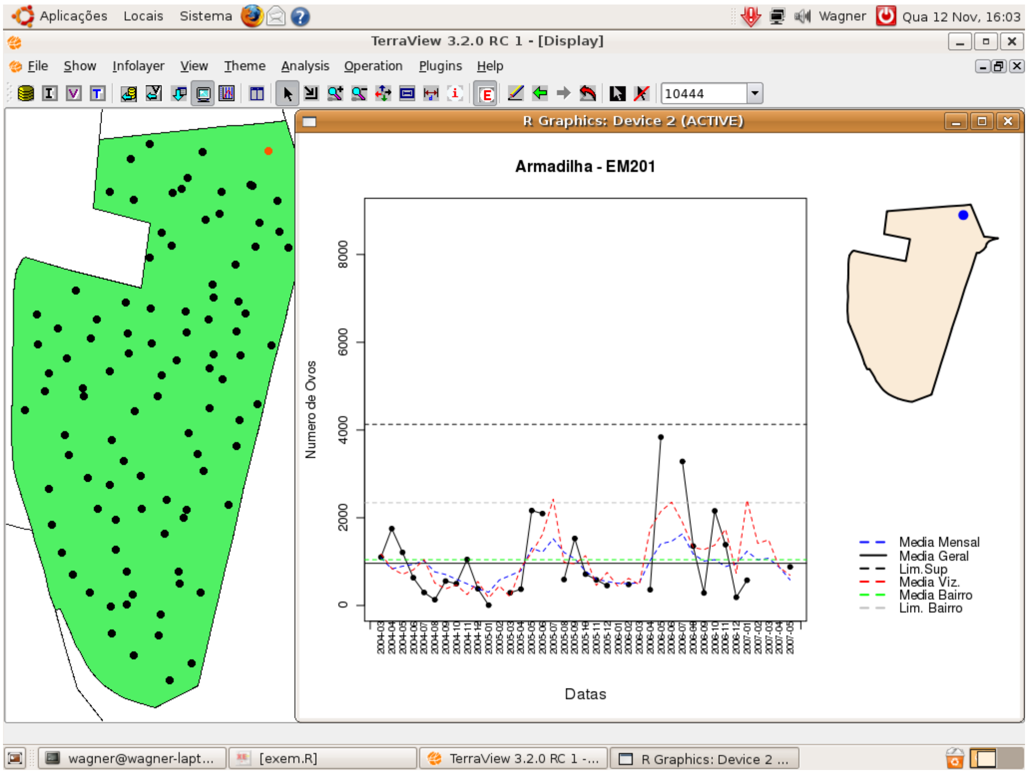


Figura 2. Análise descritiva por armadilha.

3.2. Análise por bairro

Um segundo tipo de análise, é um mapa da superfície de densidade de ovos, que permite a identificação das chamadas *zonas quentes*, buscando evidenciar regiões críticas do bairro, que devem receber prioridade de atendimento do setor de vigilância entomológica, considerando que maior número de ovos aumenta o risco da população desta área ser infectada. Vários métodos para estimar esta superfície estão implementados no RDengue, para citar alguns tem-se: modelos aditivos generalizados[Wood 2006], superfícies de tendência[Druck et al. 2004], regressão local [Cleveland 1978], e diferentes ajustes utilizando funções suavizadoras (*splines*)[Wahba 2000]. Busca-se com este tipo de análise, estimar uma superfície suave cobrindo todo o bairro. Conforme o método utilizado pode-se explorar diferentes pontos de análise. Por exemplo, utilizando a regressão local se evidência os efeitos espaciais locais tais como efeitos de vizinhança entre armadilhas. Enquanto, que a superfície de tendência evidência efeitos globais, como tendências direcionais. Cada método tem suas particularidades, e pode ser mais ou menos adequado de acordo com o objetivo da análise, e a configuração observada nos dados. Uma outra opção, é usar a transformação logarítmica, como os dados são originalmente uma contagem esta transformação é recomendada constantemente na literatura por estabilizar a variância e trazer gaussianidade, pressuposto de alguns dos métodos utilizados para estimar as superfícies. E ainda, outra opção seria gerar as superfícies por uma, duas ou todas as técnicas implementadas e compará-las, já que, este procedimento é facilitado pelas rotinas do RDengue. A Figura 3 mostra a superfície estimada por quatro métodos, para o bairro Engenho do Meio da cidade de Recife/PE. A superfície (a) traz o ajuste usando

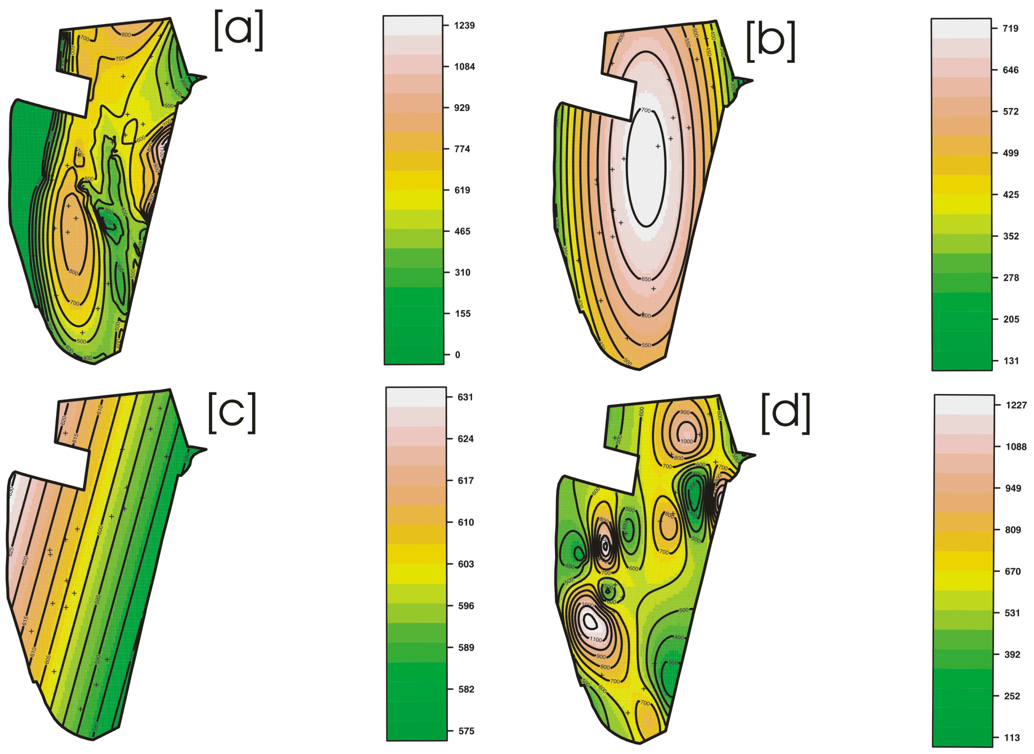


Figura 3. Exemplos de mapas de suavização.

uma Regressão Local (loess). A superfície em (b) foi ajustada utilizando o método de superfície de tendência. A superfície em (c) utiliza um Modelo Aditivo Generalizado e a superfície em (d) foi ajustada usando funções splines multiníveis. A partir dos mapas de suavização é possível investigar de forma exploratória padrões espaço-temporais gerando animações. Para isto, utiliza-se uma interpolação *pixel* a *pixel*, o que resulta em um 'filme' que mostra a evolução do fenômeno tanto no espaço como no tempo. Os complementos online deste trabalho, mostram tal filme utilizando a escala de cores topográfica, em que valores altos da densidade de ovos são representados por cores claras sendo o valor mais alto pela cor branca, no outro extremo, valores baixos são representados por tons de verde. Considera-se, que este tipo de ferramenta exploratória é muito importante nas fases iniciais de modelagem estatística, onde se busca explorar padrões dos dados que possam sugerir estratégias adequadas de modelagem.

4. Conclusões

Com a concepção de construção de ambientes para o monitoramento e vigilância entomológica, o RDengue soma-se aos objetivos do SAUDAVEL, contribuindo para o aumento da capacidade do setor de saúde no controle de doenças transmissíveis, desenvolvendo instrumentos para a prática da vigilância entomológica, incorporando aspectos ambientais, identificadores de risco e proteção, e métodos automáticos e semi-automáticos que permitem a detecção de surtos e seu acompanhamento no espaço e no tempo.
As rotinas implementadas no RDengue mostram-se satisfatórias, para gerar análises exploratórias, tendo inclusive um procedimento para a visualização espaço temporal, que evidência picos e tendências espaciais mostrando a evolução do fenômeno em toda a área, dando a possibilidade de ver os dados, procedimento nada trivial devido ao tamanho do

experimento. Também possibilita, um monitoramento da intensidade de ovos do mosquito em toda a área com base em amostras pontuais do fenômeno, sendo que podem ser definidos sinais de aviso quando, por exemplo, a intensidade de ovos capturados for muito superior ao padrão considerado normal. Através dos modelos *arima* escolhidos de forma ótima para cada uma das 564 armadilhas, pode-se fazer a previsão para *k*-passos a frente. O intuito de uma análise exploratória é dar subsídios para a busca de modelos que se adequem aos dados, fornecendo a visualização de como o fenômeno se desenvolveu tanto no espaço como no tempo.

Desenvolvimentos futuros devem incorporar outras possíveis covariáveis aos modelos que estimam as superfícies tais como, temperatura, umidade e precipitação do ar. Métodos de previsão devem considerar extensões e alternativas aos modelos *arima*, gerando possíveis cenários futuros das armadilhas com maior acurácia. A avaliação de medidas de combate ao mosquito guiada por modelos estatísticos deve ser avaliada, e incorporar a capacidade de gerar cenários sob determinados tipos de intervenção.

As tecnologias utilizadas aqui e as rotinas do RDengue devem se acoplar aos ferramentais desenvolvidos pelo projeto SAUDAVEL e disponibilizadas a orgãos responsáveis pela vigilância entomológica, contribuindo desta forma para o setor de saúde na prevenção e detecção de surtos de dengue.

O pacote RDengue é disponibilizado em *www.leg.ufpr.br/RDengue*, junto a uma breve descrição de suas funcionalidades e suas dependências. Os códigos, resultados e telas de análises utilizados aqui são disponibilizados em *www.leg.ufpr.br/papercompanions* que corresponde ao complemento online deste artigo.

Agradecimentos

Este trabalho é parcialmente financiado pelo projeto SAUDAVEL, CNPq 552044/2002-4 e desenvolvido no LEG/UFPR, em parceria com integrantes da rede SAUDAVEL da DPI/INPE, ENSP/Fiocruz, RJ e CPqAM/Fiocruz. Agradecemos a toda equipe da rede bem como aos técnicos de campo e laboratório da Secretaria de Saúde do Recife/CVA e também a população dos bairros de Engenho do Meio, Brasilia Teimosa, Morro da Conceição, Alto José do Pinho, Sítio dos Pintos, Dois Irmãos, Casa Forte e Parnamirim pela colaboração.

Referências

Akaike, H. (1974). A new look statistial model identification. *IEEE Transactions on Automatic Control AU-19*.

Andrade Neto, P. R., Ribeiro Jr., P. J., and Fook, N. D. (2005). Integration of Statistics and Geographic Information System: the R/Terralib case. *VII Simpósio Brasileiro de Geoinformática*.

Cleveland, W. (1978). Robust locally weighted regression and smoothings scatterplots. *Journal of the American Statistical Association*.

Câmara, G., Souza, R. C. M., Pedrosa, B. M., Vinhas, C., Monteiro, A. M. V., Paiva, J. A., Carvalho, M. T., and Gattass, M. (2000). Terralib: Technology in Support of GIS Innovation. *II Simpósio Brasileiro de Geoinformática*.

Druck, S., Carvalho, M. S., Câmara, G., and Monteiro, A. M. (2004). *Análise Espacial de Dados Geográficos*. Embrapa, Brasília,DF.

Hyndman, R. J. and Khandakar, Y. (2008). Automatic times series forecasting: The forecasting package for R. *Journal of Statistical Software*.

Monteiro, A. M., Carvalho, M. S., Assunção, R., Vieira, W., Ribeiro Jr, P. J., Davis Jr, C., and Regis, L. (2006). SAUDAVEL: Bridging the Gap between Research and Services in Public Health Operational Programs by Multi - Institutional Networking Development and Use of Spatial Information Technology Innovative Tools.

Montgomery, D. C. (2000). *Introduction to Statistical Quality Control*. John Wiley e Sons, 4th edition.

R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Regis, L., Monteiro, A. M., Santos, M. A. M., Silveira, J. C., Furtado, A. F., Acioli, R. V., Santos, G. M., Nakazawa, M., Carvalho, M. S., Jr, P. J. R., and Souza, W. V. (2008). Developing new approaches for detecting ans preventing Aedes aegypti population outbreaks: bases for surveilance, alert and control system. *Memórias do Instituto Oswaldo Cruz*.

Silveira, J. C., Souza, W. V., Régis, L. N., Santos, M. A., Lapa, T. M., Portugal, J. L., Braga, T. S., and Monteiro, A. M. (2004). Recife em pedaços: Geotecnologias para a Detecção e Acompanhamento em Vigilância epidemiológica. *VI Congresso brasileiro de epidemiologia*.

Tauil, P. L. (2002). Aspectos críticos do controle do dengue no brasil. *Cadernos de Saúde Pública [online]*.

Wahba, G. (2000). *(Smoothing) Splines in Nonparametric Regression*. University of Wisconsin - Technical Report, Madison, Department of Statistics.

Wood, S. N. (2006). *Generalized Additive Models: Introduction with R*. Chapman and Hall, Boca Raton.

Using Serious Game Techniques to Simulate Emergency Situations

Marcelo G. Metello, Marco A. Casanova, Marcelo T. M. Carvalho

Departamento de Informática
Pontifícia Universidade Católica (PUC) – Rio de Janeiro, RJ – Brazil
{metello,casanova}@inf.puc-rio.br, tilio@tecgraf.puc-rio.br

Abstract. *This paper presents a simulation engine we implemented to support an interactive emergency simulation game. We first discuss various issues on building serious (non-entertainment) games involving the simulation of real life situations. The discussion also considers the use of geographical information systems and dynamic modeling techniques. Then, we present the architecture of the simulation engine and discuss the main aspects regarding its application to emergency plans. Finally, we briefly describe a prototype built to validate the proposed architecture and the requirements raised in the discussion.*

Resumo. *Neste trabalho apresentamos um simulador implementado para atender as necessidades de jogos interativos simulando situações de emergência. Inicialmente, apresentamos os problemas inerentes ao desenvolvimento de jogos sérios (sem propósito de entretenimento) que envolvem a simulação de situações do mundo real. A discussão também considera a adoção de sistemas de informação geográfica e de técnicas de modelagem dinâmica. Em seguida, apresentamos a arquitetura de um simulador e discutimos os aspectos principais levantados pela sua aplicação no contexto de planos de emergência. Por fim, descrevemos brevemente um protótipo construído para validar a arquitetura proposta e os requisitos levantados.*

1. Introduction

The first computer games were created in the early 1960's. Since then the computer game industry grew into a large segment that now plays a relevant role in the evolution of several areas of Computer Science, such as human-computer interfaces, computer graphics, artificial intelligence and, more recently, computer networks [Smed et al 2002].

The widespread adoption of computer games for entertainment purposes, the continuous decrease of hardware cost and the success in military simulations made gaming technologies attractive to some "serious" industries such as medicine, architecture, education, city planning, and government applications [Smith 2007]. The term *serious games* [Susi, Johannesson and Backlund 2007] has been used to denote games used for such non-entertainment purposes. The application of gaming technologies to these areas presents some peculiar challenges, since their requirements can be quite different from those of the entertainment industry. Usually, serious games need to work on models which reproduce certain aspects of reality. On the other hand, entertainment games have much more freedom to create and modify their own reality, which can be quite convenient, especially when the developers face technical limitations. Even though entertainment games may require realistic audiovisual player experience, they do not need to reproduce realistic situations.

Apperley (2006) classifies video games into four genres according mainly to the characteristics of the player interaction. Specifically, the *simulation* genre better defines the applications we focus in this paper. Still according to the author, simulation games can be further analyzed with respect to their degree of realism, which is the main aspect that characterizes serious games. We will therefore use the term *serious simulation games* to denote serious games that belong to the simulation genre.

The fact that serious games may require realistic simulations justifies the effort to integrate gaming techniques with traditional *geospatial dynamic models*. The purpose of geospatial dynamic models is to describe processes that have some important spatial aspect. Such processes may include natural phenomena and human action on Earth. Examples of geospatial dynamic models are extensively found in the literature related to fields such as hydrology, climate changes, land use, population dynamics and many others. Such models improve our understanding of dynamic phenomena as they make it explicit the causal relationships between the elements involved.

Since geospatial dynamic models attempt to describe real phenomena, they help meet the requirements of serious games for realism. Hence, the effort to make dynamic modeling engines interoperate with simulation game engines is perfectly justifiable. Ideally, the player interaction capabilities of computer games and realistic dynamic modeling techniques should be integrated in a complementary way.

This paper illustrates the combination of these two approaches to simulate an emergency response activity. An emergency situation occurs when an incident can cause damage to human health and the environment. Once it occurs, the best approach to control it is to respond quickly and in a well organized manner. Testing the performance of an emergency response team is mandatory to ensure minimum impact of the incident. Testing usually takes the form of field exercises, but simulation has also been successfully used. We discuss the advantages of adding simulation capability to an emergency information management system, such as InfoPAE [Carvalho et al. 2001], a system used to manage emergency situations at Petrobras, the Brazilian oil company.

The paper is organized as follows. Section 2 lists the requirements for serious simulation games. Section 3 describes a simulation engine that integrates traditional dynamic models, and discusses issues related to player interaction in simulation games with geospatial aspects. Section 4 presents the emergency simulation game. Finally, Section 5 contains the conclusions and enumerates future work.

2. Requirements for Serious Simulation Games

This section lists some of the requirements for serious simulation games that are not fundamental to other classes of computer games.

2.1. Realism

There should be a minimum acceptable degree of realism in serious simulation games. This suggests that the simulations involved in this kind of game should run on data that represents real objects, which is precisely the difference between graphical and geographical data.

In many cases, the ability to access data in existing systems, such as geographical information systems (GIS), will work as a subrequisite to realism. Indeed, in order to create a realistic simulation, the system has to work with real data, which is likely to be stored in existing information systems and databases.

2.2. Time Flow Control

The ability to stop, accelerate and go back in time can be quite important when designing serious simulation games. It is not difficult to imagine situations where it would be desirable to pause or to replay a simulation multiple times from a given point. In other situations, it may be desirable to accelerate the time flow, especially in periods without much player activity.

This requisite says that serious simulation games may require much more control over the time flow than entertainment games. Some serious simulation games may require that entire simulations be recorded for future replays.

2.3. Learning Support

Simulation games devoted to training require player evaluation as part of the learning process [Borodzicz and van Haperen 2002]. In many cases, this requirement means that the system should be able to play back a simulation for evaluation purposes, which involves saving simulation data.

Apart from simply saving simulation data, the system may be required to trace back player decisions for payoff evaluation [Zagal et al 2006]. In this case, the underlying simulation system should be aware of the player action model.

There are also cases where organizations will have predefined procedures and protocols. When this happens, it may be necessary to match the sequence of player actions to these predefined procedures to check whether each player acted as expected by the organization. Going one step further, the simulations can be used to evaluate the effectiveness of the predefined procedures, detect their flaws and help with their evolution [Smith 2004].

It should be noted that the requirement of learning support is not limited to individual learning. The evaluation of the effectiveness of predefined procedures represents some kind of collective or institutional learning.

3. The Simulation Engine

This section provides a high level description of the architecture of the implemented simulation engine, and discusses some interesting issues related to it.

In the context of this work, a simulation refers to a set of elements and events evolving in a spatio-temporal representation of the world. That is, the world state is represented by spatial and non-spatial data, and the state changes as time flows. The changes in the state of the world are affected by input given by the players, which in part justifies why the simulation engine may also be considered a game engine.

3.1. Architecture

Figure 1 illustrates the high level architecture of the system. The simulation engine is responsible for continuously updating the state of the world as the simulation time flows. The renderer is responsible for generating visual input for the players, and is not considered as part of the simulation engine. This makes it possible to implement a multiplayer game with different renderers for different players, and to generate different views for each player, according to which information they are allowed to see.

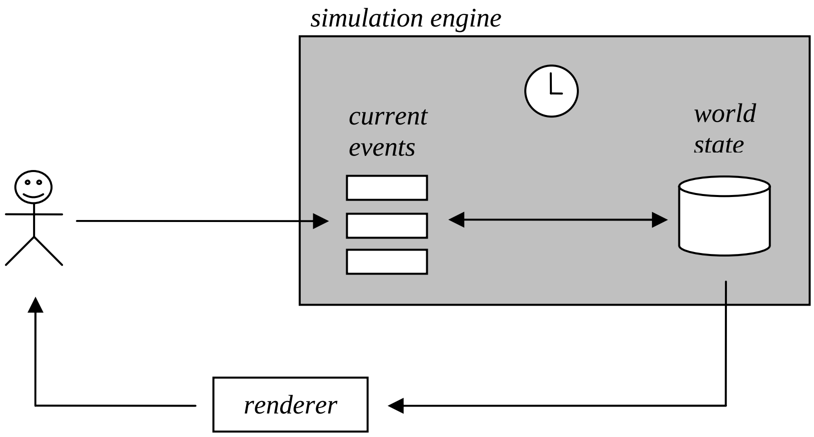


Figure 1. The high level architecture of the system.

All changes in the world state are carried out by special entities called *events*. There are two types of events: *player events* and *system events*. Player events represent player activity in the world. The way player input alters the state of the world is always defined by player events. On the other hand, system events represent all activity that is not directly caused by any player.

Events have duration. Therefore each event has a start time and a finish time. As an example, the event of moving a boat to a specified location at sea would start at the moment the order to move is given and finish at the time the boat gets to the specified location.

The lifecycle of a player event starts when the player issues a command to execute it. The player event then starts its activities by generating a series of instructions. Each instruction is basically a piece of code that somehow alters the state of the world in the simulation. The instructions are always instantaneous with respect to simulation time. Each instruction is assigned a timestamp, which will provide a means of ordering instructions originated from different concurrent events. In the case of the moving boat player event, each instruction would change the location of the boat to the next point in its trajectory to its destination. Note that, since each player event generates its instructions, they are responsible for determining the granularity of their execution. It is also possible for a player to have multiple events running at the same time.

There are three possible outcomes of the execution of a player event. It may *finish* its execution successfully, it may be *cancelled* or it may *fail*. A player event can be cancelled in three situations. The player may issue a command to cancel one of his running player events. If the player issues a command to execute one player event that conflicts with another player event, one of them is cancelled by the system. As an example, if the player has ordered a boat to go to location *L* and, before the boat gets to *L*, he orders the same boat to go to another location *L'*, the event of moving the boat to *L* is cancelled. As the last possibility, when the simulation finishes, all running events are cancelled. Fig 2 illustrates the state diagram for player events.

The event model implemented by the simulation engine should be general enough to implement various kinds of serious games. One possible indication of this generality is that it is very similar to John Sowa's process model [Sowa 2000].

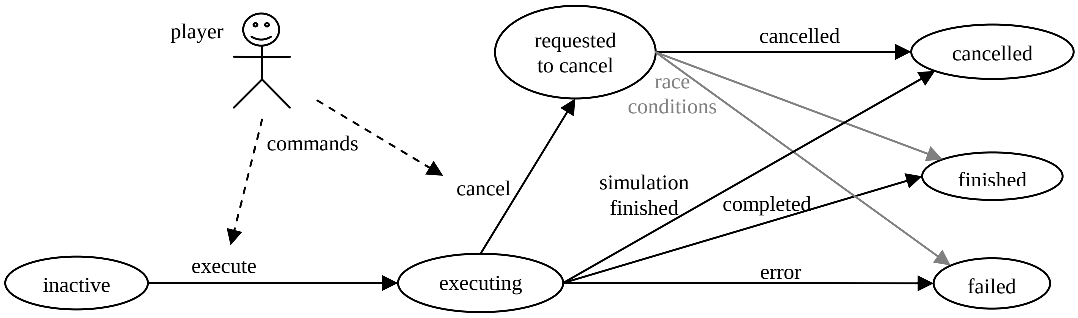


Figure 2. The state diagram of player events.

3.2. Issues related to dynamic modeling

Dynamic modeling frameworks, such as PCRaster [van Deursen 1995], Swarm [Minar et al 1996] and Nested-CA [Carneiro 2006] usually provide or use a specific language to describe dynamic models and an engine to run these models on real data, usually provided by GIS.

Since the event model used to implement the simulation engine is very generic, it is usually possible to implement most dynamic models as system events. In fact, it should be possible to emulate any dynamic model engine inside one system event. In this case, the event would issue an instruction every time the dynamic model engine should alter the state of the world. In fact, this is how the simulation engine proposed in this paper integrates traditional dynamic modeling and games.

Another issue is related to running time. Simulation engines and frameworks designed for dynamic modeling seldom show any concern for running the models in real time, which is a requirement for interactive games. The term *real time* is used here to denote that the time flow in the execution of a temporal model should be synchronized with the real time flow for better user experience.

The technology developed by the gaming industry is focused on displaying data in real time for highly interactive gaming. If we go through the process of game development, there is a continuous attention on performance requirements in almost every part of the code. Performance improvements are tried everywhere to keep an acceptable frame rate for better user experience. This objective is always present throughout the game development process.

If the dynamic models and the spatial data are not too complex, the real time requirement can be implemented simply by adding a time controller to an existing dynamic modeling engine. However, since GIS are known for their heavy datasets, techniques for optimizing spatio-temporal data manipulation may be necessary [Wolfson et al 1998, Siebeck 2004].

3.3. Issues related to GIS

GIS have been traditionally more concerned with displaying spatial data than temporal data. Only recently the major GIS, other than military and flight simulators, started to take into account time and user experience, much in the same way as for the gaming industry. The increasing popularity of training games certainly contributes to this change.

GIS are known for their heavy spatial datasets. This is certainly one of the main reasons why it is difficult to display animated GIS data at a minimum acceptable frame rate. Metello et al. (2007) show how the fast computer graphics techniques used in gaming may be used to display GIS data at higher frame rates.

3.4. Issues related to multithreading

Most information systems and GIS follow a single-thread paradigm in their software architecture. The paradigm is characterized by a synchronous model of work where the user issues a command and waits for the response of the system before he can issue more commands. This type of architecture is clearly not adequate for interactive simulations, since the simulation should be run in real time and cannot afford waiting for user input. The simulation must continue even if the user does not issue any command. Besides, the system must also support multiple players, which is another argument against the single-thread paradigm.

In more dynamic applications, such as entertainment action games, the system executes a process continuously, regardless of commands input by the user in an asynchronous way. The differences in the execution flow of both kinds of architectures are illustrated in Fig. 3. In the workflow on the right, there are two processes running asynchronously. In the simulation engine, the interaction between these two processes is handled simply by queuing *player events* when they are created by a player. Of course, the queue must be thread-safe.

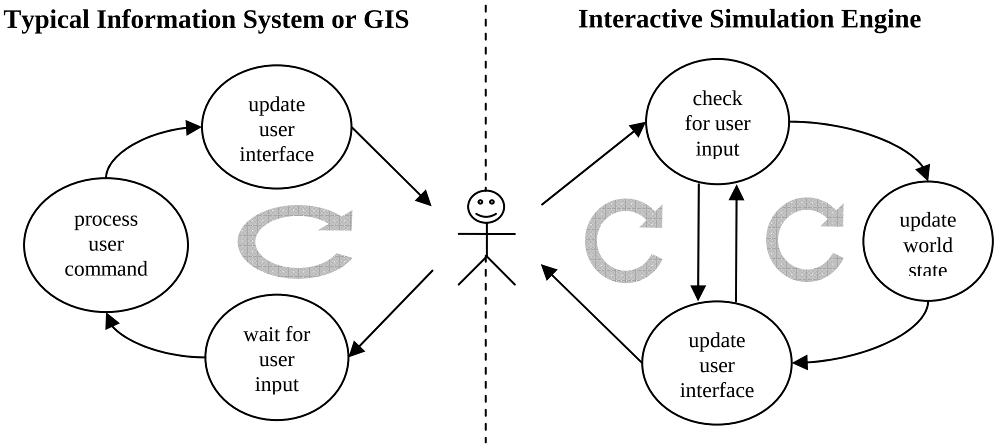


Figure 3. The kind of architecture required for interactive simulations (right)

4. The Interactive Emergency Simulation Game

In order to validate the proposed architecture, a game was implemented with the objective of simulating some specific emergency situations. In the context of this paper, an emergency is an incident, like an oil spill, that requires response action to be controlled and to mitigate effects, as loss of life or damage to property and natural resources.

Preventing the incident is always the best for avoiding damage to human health and the environment. However, once an emergency occurs, the best approach to control it is to respond quickly and in a well organized manner. This will happen if response strategies have been planned ahead of time. One of the elements for this planning is the emergency plan, which comprises a set of instructions that outline the steps that should be taken before, during, and after an emergency. The emergency plan is based on a risk assessment that looks at all the possibilities of what could go wrong. To assist and support a better response action, the plan contains, besides the set of instructions, the response strategies, a list of contacts and personnel, a material resource list, and refers to a vast documentation.

After the plan is developed, it is important to test it to check whether it works as expected and to train and evaluate the operational readiness of responders. Testing usually takes the form of an exercise or drill, what can be very time consuming and expensive to organize periodically. Another point to consider is the difficulty for representing detailed and realistic situations required to effectively test the emergency plan. Use of simulation games in these cases can be helpful.

This section further discusses the motivation outlined above and the implementation of a simulator.

4.1. Representation versus simulation: limitations of planning response ahead of time

In [Frasca 2003], the author discusses two different approaches for modeling knowledge about dynamic phenomena: representation and simulation. According to the author, the main difference between both forms is that simulation attempts to model the behavior of the elements involved in the phenomenon while representation is limited to retaining the perceptual characteristics of it. To make it clear, the author gives the example of a plane landing procedure. A representation of a specific landing could be a film where an observer would be incapable of interfering. On the other hand, a flight simulator would allow the player to modify the behavior of the system in a way that simulates the real plane. This flexibility is only possible due to the simulation characteristic of modeling the behavior of the elements independently of any specific scenario.

An emergency plan takes, traditionally, a more representational form. It contains response strategies planned for different type of scenarios, but it cannot tell whether the plans are well suited for all cases. The set of instructions contained in an emergency plan consists basically of workflows of actions. They do not take into consideration the preconditions or durations of each action. Moreover, they do not take into consideration the specific spatial characteristics of all scenarios. For example, a plan can describe the action of sending two boats to intercept an oil spot. However, it may not be possible to do that before the oil reaches the coast in some specific conditions. If emergency managers were able to simulate the whole process in a more realistic way, it would certainly make the emergency plans more reliable.

In order to meet all these needs, a simulator was developed with the purpose of simulating emergency scenarios. Some of the main advantages of building an emergency simulator and integrating it with an emergency management system include:

- Simulations help finding flaws in emergency plans
- Testing whether available emergency response resources configuration are enough to handle any scenario requirements
- Simulation games provide training that help improve personnel performance
- Computer simulation cost are significantly lower than functional or full scale exercises

The first scenario considered to test the simulator was the spill of a considerable volume of oil into the ocean. This scenario was chosen because it involves elements of dynamic modeling and user actions that interfere with each other.

4.2. The Oil Spill Scenario

In a typical oil spill scenario, the initial goals are to attempt to control the leak at the source of the spill, and to limit the propagation of the floating oil as much as possible. These goals are typically achieved by using containment, recovery and clean-up strategies, which are implemented through specific operational procedures. These procedures, however, depend on the characteristics of the scenario, such as:

- Type of oil and its characteristics
- Local of the source of the leak and its nearest landmark
- Estimation of the amount of oil spilled
- Weather and sea conditions
- Characteristics of the shoreline that may be affected

4.3. The Emergency Simulator

The first simulation implemented considers the oil spill scenario after the leak has stopped and focuses on oil contention in water bodies. Clean-up operations for oil that reached the coast were also not considered. The goal of this simulation is to test emergency plans for leaked oil containment. The simulator is expected to uncover possible flaws in the plans as well as to help planning equipment installation and location.

The initial conditions of the simulation are specified in a document which is read by the system. This document includes the location, amount and type of leaked oil, maps of the affected area, the location of all available equipment and weather conditions. The dynamic elements of the simulation are described next.

The leaked oil is represented by a hexagonal cell space, where each cell stores the amount of oil in it. The movement of oil is modeled using a cellular automaton, which considers not only the state of each cell but also the weather conditions, such as the wind direction and speed, for example. Currently, weather conditions are globally defined.

The oil movement model must also consider elements that will act as obstacles, such as shorelines and barriers used to contain the oil spilled. Both are represented as polylines in the simulation and have similar treatment. Each cell that intersects either a barrier or a coast line is considered an obstacle cell.

When the oil reaches the coast, its movement must take into consideration the type of the coast. For example, sand coasts absorb oil in a much greater amount than rocky coasts. It will be helpful to use a detailed map of coast types for all locations that can be reached by the oil spill. Likewise, different types of barriers have different absorption and containment capabilities.

Each cell in the cellular space must contain information as whether it represents an obstacle or not. Note that this state may change during the simulation when barriers are used. Since different obstacles have different containment and absorption characteristics, the obstacle cells must also keep this information. All the information about obstacle cells is of course used as input to the cellular automaton as well.

The main dynamic elements in the simulation other than the oil itself are the boats used to support operational response to the emergency situation. Boats are responsible for placing barriers to contain oil. They may also carry some oil recovery equipment, such as pumps and skimmers. The initial location of the boats is defined in the document that describes the initial conditions for the simulation.

The movement of the boats is simulated by taking into account their speed and cargo capacities, as well as weather conditions, such as wind, sea currents and tide. During the simulation, players guide the boats through way-points. They may place way-points wherever they want and send the boats to any of them. Of course, boats may also encounter obstacles such as islands. In this case, they just stop and wait for further instructions.

The placement of barriers is an operation executed by two boats. Each boat holds one end of the barrier. The idea is that, when the oil spot passes between the boats, it gets blocked by the barrier. After that, the recovery operation starts. The command to place a barrier needs some parameters, such as which barrier should be used, the angle at which it should be put, the distance the boats should keep from each other and the curvature that should be kept. In order to execute the action of placing the barrier, some preconditions must be met. The two boats must not be too far away, the barrier must be long enough for the distance and curvature given as parameters, and the weather conditions must not prevent the placement of the barrier.

5. Conclusion and Future Work

The architecture used in the simulation engine proved to be a way of integrating traditional dynamic modeling with interactive computer games. This is crucial to give proper realism to a simulation, as an emergency response simulation.

Although the simulator described helps training and evaluating emergency plans, the analysis of the performance of the players is still a manual process. Future versions should consider the possibility of integrating the flow of player actions with the predefined workflows in emergency plans in an automatic way.

Another issue of interest is to integrate the simulator with an emergency information management system, which usually holds a detailed database of emergency plans for different types of scenarios, but claims for representing detailed and realistic situations where to test emergency plans. InfoPAE [Carvalho et al. 2001] provides an example of such systems. InfoPAE was designed to manage emergency situations at Petrobras, the Brazilian oil company. Besides the emergency plans, the database also provides detailed data about procedures, material resources, facilities, available equipment and geographical data.

Using the simulator with such systems will provide a good environment to test the emergency management circular process through which managers prepare for emergencies, respond to them, recover from them and mitigate their effects, and prevent future emergencies from occurring. In this sense, simulation plays a critical role to assess, improve performance and prepare for future emergencies.

The scenarios used to test out prototype were derived from the InfoPAE databases. It might be interesting to create a script language for defining different emergency scenarios, even if they are hypothetical. This could help improving the training process by generating sequences of scenarios with increasing difficulty levels, just like in entertainment games.

References

Apperley, T. (2006) "Genre and game studies: Toward a critical approach to video game genres". Simulation & Gaming, 37(1), 6-23

Borodzicz, E. and van Haperen, K. (2002) "Individual and Group Learning in Crisis Simulations". Journal of Contingencies and Crisis Management 10 (3), 139–147 doi:10.1111/1468-5973.00190

Carneiro, T. (2006) "Nested-CA: A Foundation for Multiscale Modelling of Land Use and Land Cover Change". Doctorate Thesis from the Post Graduation Course in Applied Computer Science, INPE - Sao Jose dos Campos, Brazil

Carvalho, M.T.; Freire, J.; Casanova, M.A. (2001) “The Architecture of an Emergency Plan Deployment System”. Proc. III Workshop Brasileiro de GeoInformática, Rio de Janeiro, Brasil, Oct. 2001.

Frasca, G. (2003) "Simulation versus Narrative: Introduction to Ludology". In: Wolf & Perron (Eds.) The Video Game Theory Reader. Routledge.

Metello, M, et al. (2007) “Continuous Interaction with TDK Improving the User Experience in Terralib”. Proc. IX Brazilian Symposium on GeoInformatics.

Minar, N., Burkhart, R., Langton, C. and Askenazi, M. (1996). “The Swarm Simulation System: A Toolkit for Building Multi-agent Simulations”. Working Paper 96-06-042, Santa Fe Institute, Santa Fe.

Siebeck, J. (2004) “Concepts for the Representation, Storage, and Retrieval of Spatio-Temporal Objects in 3D/4D Geo-Information-Systems”. Rheinische Friedrich-Wilhelms-Universitat Bonn, Diss.

Smed, J., Kaukoranta, T., Hakonen, H. (2002) “Aspects of Networking in Multiplayer Computer Games”. The Electronic Library, Volume 20, Number 2, 2002, pp. 87-97(11)

Smith, D. (2004) “For Whom the Bell Tolls: Imagining Accidents and the Development of crisis Simulation in Organizations”, Simulation & Gaming, 35(3): 347-362

Smith, R. (2007) “Game Impact Theory: Five Forces That Are Driving the Adoption of Game Technologies within Multiple Established Industries”. Games and Society Yearbook.

Sowa, J. (2000) “Knowledge Representation: Logical, Philosophical, and Computational Foundations”. Brook/Cole, a division of Thomsom Learning: Pacific Grove, CA.

Susi, T., Johannesson, M., Backlund, P. (2007) “Serious Games – An Overview”. Technical Report HS-IKI-TR-07-001, School of Humanities and Informatics, University of Skövde, Sweden

Van Deursen, W.P.A. (1995) “Geographical Information Systems and Dynamic Models”. Ph.D. thesis, Utrecht University, NGS Publication 190, 198 pp. Electronically available through www.carthago.nl

Wolfson, O., Xu, B., Jiang, L., Chamberlain, S. (1998) “Moving Objects Databases: Issues and Solutions”. SSDBM, p. 111, 10th International Conference on Scientific and Statistical Database Management

Zagal, J.P., Rick, J., Hsi, I. (2006) “Collaborative games: Lessons learned from board games”, Simulation & Gaming, 37(1), 24-40

Simulando padrões de incêndios no Parque Nacional das Emas, Estado de Goiás, Brasil

Rodolfo M. Almeida[1], **Elbert E. N. Macau**[2], **Helena França**[3], **Fernando M. Ramos**[2], **Tiago G. de Senna Carneiro**[4]

[1]Programa de pós-graduação em Computação Aplicada, INPE

[2]Laboratório Associado de Computação e Matemática Aplicada, INPE

[3]Divisão de Geofísica Espacial, INPE

[4]Departamento de Computação, UFOP

{rodolfo, elbert, fernando}@lac.inpe.br, helena@dge.inpe.br, tiago@iceb.ufop.br

Resumo. *Apresentamos neste trabalho um modelo probabilístico e espacialmente explícito para simular padrões de incêndios ocorridos no Parque Nacional das Emas, Estado de Goiás, Brasil. Aplicamos conceitos da teoria de percolação para representar os fenômenos de ignição e propagação. O fogo se propaga de uma célula queimando para qualquer uma das oito células vizinhas como um evento estocástico independente com probabilidade I, onde I pode variar de 0 a 1. Uma célula com material combustível permanece em seu estado caso nela não ocorra a ignição. Uma vez levada à ignição, seu estado será, nas duas iterações seguintes, queimando e queimado. A extinção do incêndio ocorre se novas células não forem levadas à ignição em um dada iteração. Efeitos do acúmulo de combustível são incluídos no modelo mediante escolhas de distintos valores de probabilidade de propagação. Os efeitos do vento são introduzidos com um incremento na probabilidade de propagação para vizinhos adjacentes situados na direção do vento.*

1. Introdução

O fogo sempre foi um elemento fascinante para o homem, tornando-se com o evoluir dos tempos, num dos elementos da natureza mais utilizados no seu cotidiano. No entanto, quando o fogo ocorre sob condições descontroladas, pode se torna perigoso e destruidor. Um incêndio florestal é caracterizado pela presença indesejada do fogo que se alastra em condições fora de controle consumindo a vegetação. Incêndios florestais indesejáveis são um dos maiores desastres naturais que ameaçam diversas regiões do mundo. Cada ano, milhares de hectares de áreas cobertas por vegetação são consumidas pelo fogo, pondo em risco o meio-ambiente e a vida do homem.

O interesse em se modelar o comportamento do fogo em incêndios florestais tem sido o objetivo de muitos grupos de pesquisa em várias partes do mundo nas últimas décadas. Nesse nível de abstração destacam-se os *modelos de propagação do fogo*, que simulam o avanço do fogo sobre a superfície por um conjunto de regras ou equações que levam em conta variáveis relacionadas com a vegetação, topografia e condições atmosféricas [15]. Os modelos de propagação do fogo podem ser tanto originados das leis que governam a mecânica dos fluidos, combustão e transferência de calor quanto de leis empíricas obtidas a partir de informações extraídas de dados experimentais ou de incêndios históricos.

Neste trabalho vamos nos ater a uso de um modelo empírico que utilizam técnicas de *percolação* e *autômatos celulares*. Modelos baseados em autômatos celulares - também chamados *modelos espacialmente explícitos* [2] - descrevem a paisagem em subunidades de área, chamadas *células*. Cada célula possui uma localização, uma vizinhança e a ela são incorporados atributos correspondentes às características ambientais. Cada célula possui um conjunto finito de estados e um conjundo de regras determinam a transição entre eles, levando em conta o estado da própria célula, das células vizinhas e fatores ambientais [13, 12, 1].

Na simulação da propagação de incêndios florestais por *percolação*, o fogo se propaga de uma célula para suas vizinhas de acordo com uma probabilidade específica, que depende de condições ambientais. Esta probabilidade é ajustada a dados de experimentos ou de históricos de incêndios. Modelos de percolação têm sido amplamente utilizados para representar o comportamento do fogo [6, 10, 14, 7]. Este trabalho visa aplicar os conceitos de teoria de percolação na simulação de cenários de propagação do fogo em paisagens naturais, com o ojetivo de obter as proporções atingidas por um incêndio em função de determinadas condições ambientais. Tais resultados se configuram como uma informação muito importante para orgãos responsáveis pelo manejo do fogo em unidades de conservação. A região-alvo de aplicação do modelo é o Parque Nacional das Emas, situado no Estado de Goiás - Brasil.

Nosso artigo é dividido em quatro partes. Na primeira, discutimos a respeito do comportamento do fogo em incêndios no Parque Nacional das Emas e sobre os dados ambientais utilizados nesse trabalho. Na segunda parte, apresentamos os conceitos envolvidos na aplicação de teoria da percolação na modelagem do comportamento do fogo e descrevemos o modelo de propagação do fogo que estamos aplicando na região-alvo. Na terceira parte experimentos de simulações são realizados visando prever o cenário de propagação de alguns incêndios históricos. Finalmente, efetuamos as conclusões e direcionamentos futuros.

2. O fogo no Parque Nacional das Emas

Criado em 1961, o Parque Nacional das Emas (PNE) possui uma área de um pouco mais que 131.000 hectares e está localizado no extremo sudoeste do Estado de Goiás, próximo às divisas com o Mato Grosso e Mato Grosso do Sul, entre as latitudes 17^{o}51’ e 18^{o}21’ S e longitudes 52^{o}43’ e 53^{o}07’ W, conforme mostra a Figura 1.

O comportamento do fogo em incêndios florestais resulta do efeito conjunto de fatores como vegetação, condições climáticas e topografia [16]. A estrutura da vegetação do Parque - quando enquadrada na descrição de tipos fisionômicos gerais como formações florestais, savânicas e campestre - indica a predominância de fisionomias abertas (savânicas e campestre) [17], caracterizando assim um ambiente para a ocorrência de incêndios de superfície, onde predominam a queima com chamas da vegetação seca acumulada sobre a superfície. Como grande parte do Parque está situada no topo de uma chapada, onde altitudes variam entre 800 m e 890 m [17], a topografia influencia pouco no comportamento do fogo. Dentre os demais fatores destacam-se como mais importantes a

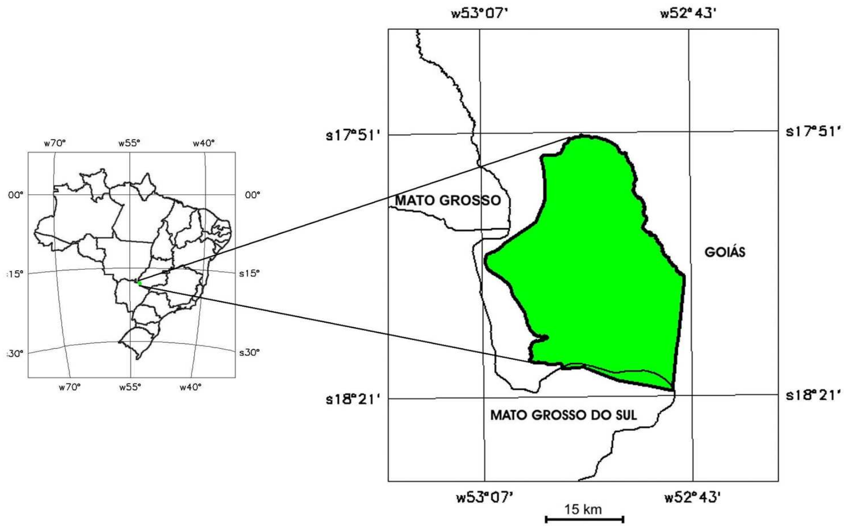


Figura 1. Localização do Parque Nacional das Emas.

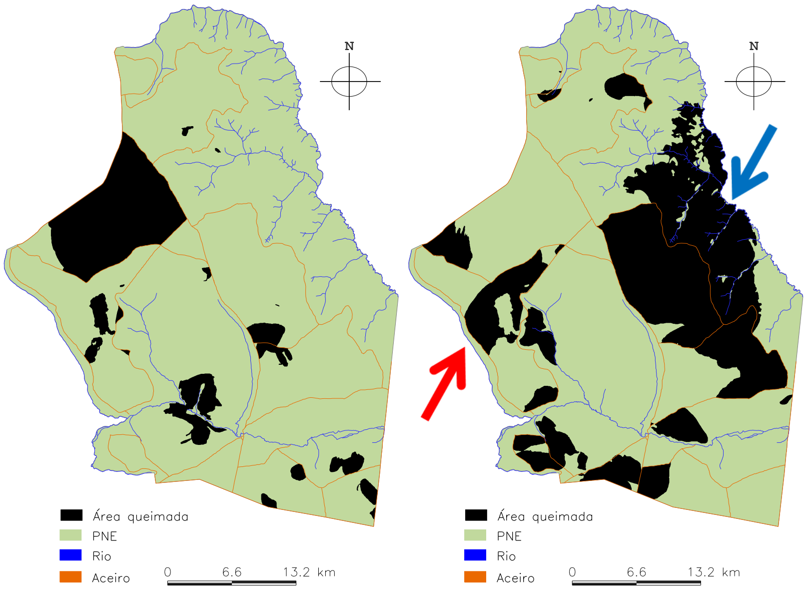
velocidade e direção do vento, e a distribuição de combustível (acúmulo e continuidade) sobre a superfície.

A continuidade do combustível pode ser interrompida por causas naturais, como a presença de rios, ou por causas artificiais, como os aceiros[1] e estradas do interior do Parque, que atuam como barreiras ao avanço do fogo. As Figuras 2(a) 2(b) mostram, respectivamente, os padrões finais dos incêndios ocorridos nos períodos de Junho de 2001 a Maio de 2002 e de Junho de 2002 a Maio de 2003. Percebe-se nas imagens que os aceiros (linhas marrons) e rios (linhas azuis) atuam como barreiras na propagação dos incêndios (áreas pretas), contendo-os ou induzindo-os a contornar os obstáculos encontrados. A seta vermelha na Figura 2(b) indica um incêndio que ultrapassou o aceiro devido à falha na manutenção dos mesmos.

O recente regime de incêndios ocorridos no PNE, com predominância de incêndios naturais iniciados por raios e quase total ausência de incêndios antrópicos, criou um verdadeiro mosaico na cobertura vegetal do Parque [9]. Com a ausência dos grandes incêndios que deixavam quase a totalidade do Parque homogênea quanto ao estado de desenvolvimento da vegetação e biomassa seca acumulada, os incêndios naturais de tamanhos, datas e localizações variadas deram origem a áreas com diferentes densidades de biomassa e fases fenológicas. Áreas recém queimadas não atingem densidade de biomassa suficiente para propragação do fogo, agindo como barreiras para incêndios vizinhos. A seta vermelha na Figura 2(b) indica um incêndio ocorrido em novembro de 2002 que contornou uma área queimada em março de 2002 (ver Figura 2(a)).

O mosaico da vegetação pode ser mapeado em função do número de anos sem queima e essa informação é muito importante para prever o risco e padrões de queimadas futuras. Assim, áreas que não foram atingidas pelo fogo há mais tempo acumulam maior

[1] O aceiro é uma técnica para prevenção de incêndios onde regiões de descontinuidades são estabelecidas na vegetação visando evitar a propagação do fogo. A malha de aceiros do PNE inclui também as estradas no seu interior e totaliza 348 km, dividindo o parque em 20 blocos. Manutenções periódicas devem ser executadas visando manter a descontinuidade de combustível acumulado na superfície.



(a) Junho de 2001 a Maio de 2002 (b) Junho de 2002 a Maio de 2003

Figura 2. Incêndios identificados para os períodos (a) Junho de 2001 a Maio de 2002 e (b) Junho de 2002 e Maio de 2003. Em (b) a seta azul indica um incêndio que ultrapassou os aceiros e a seta vermelha indica um incêndio que não queimou uma área com pouco acúmulo de combustível, a qual foi originalmente queimada em (a).

quantidade de biomassa combustível, tornando-se portanto, mais suscetíveis ao fogo. Inversamente, as regiões mais recentemente atingidas pelo fogo têm menor risco de queimar novamente.

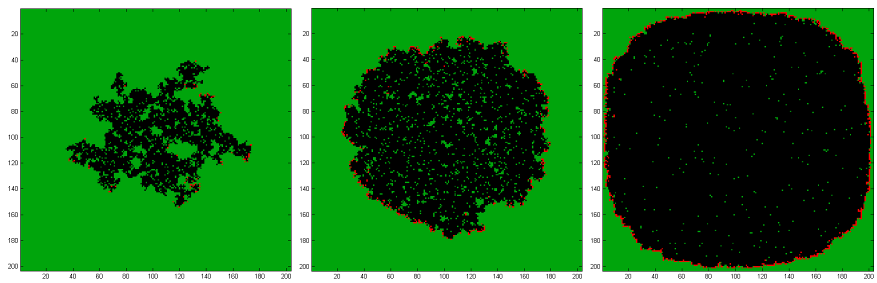
3. Modelo de propagação do fogo

O modelo aqui proposto trata a paisagem como um um reticulado bidimensional de células quadradas com dimensões 30 metros x 30 metros (900 metros quadrados de área). O atributo ambiental de cada célula corresponde ao tipo de cobertura, podendo ser célula combustível (vegetação) ou célula não-combustível (aceiro e rio). Células com material não-combustível não sofrem transição de estado. Uma célula com material combustível mudará seu estado caso seja levada à ignição. Consideramos que uma célula queima em um simples passo de tempo. Uma vez ocorrendo a ignição, o estado da célula combustível nas duas iterações seguintes serão, na respectiva seqüência, *queimando* e *queimado*. A Figura 3 ilustra as transições de estados assumidos em função das características ambientais da célula.

O conceito de *percolação de ligação simples* é aqui utilizado para representar o fenômeno de ignição e propagação. O fogo propaga de uma célula que está *queimando* para uma célula combustível vizinha como um evento estocástico independente, com probabilidade I, onde I pode variar entre 0 e 1. A vizinhança de cada célula compreende as oito células no seu entorno e a cada iteração do modelo as possíveis ligações são efetivadas ou não segundo a probabilidade I. A propagação do fogo se extingue em uma dada iteração se novas células não forem levadas à ignição em um dada iteração. No modelo, baixos valores para a probabilidade de propagação I produzem padrões de queima

Figura 3. Transições de estado das células em função do atributo ambiental. Uma célula não-combustível, que pode ser aceiro ou rio, sempre permanece em seu estado. Uma célula combustível permanece em seu estado caso ela não seja levada à ignição. Uma vez ocorrendo a ignição no passo de tempo i, o estado da célula combustível no passo de tempo $i+1$ será queimando (cor vermelha) e depois em $i+2$ será queimado (cor preta).

"dendríticos", onde a propagação flui de forma suave e lenta e o fogo aos poucos se extingue, enquanto altos valores de probabilidade produzem padrões sólidos, similares a incêndios que se propagam rapidamente. Esses efeitos podem ser vistos na Figura 4. Acima de um valor de probabilidade crítica I_c, chamada de *limiar de percolação*, o fogo "percola" livremente pela superfície [18]. O valor utilizado aqui para esse limiar de percolação foi extraído da literatura [10] e equivale a $I_c = 0,25$.

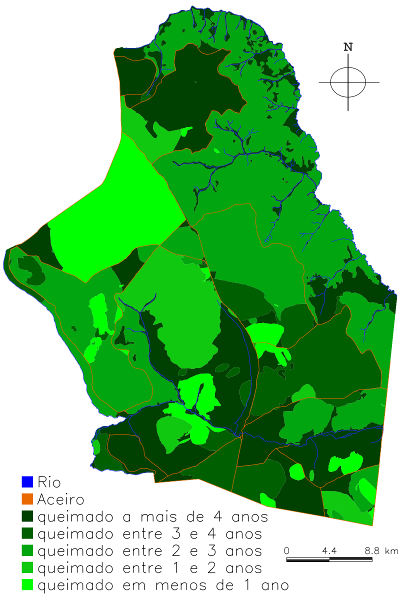


(a) $I = 0,25$ e $t = 100$ (b) $I = 0,30$ e $t = 100$ (c) $I = 0,45$ e $t = 100$

Figura 4. Padrões produzidos após 100 passos de tempo para diferentes valores de I em uma paisegem homogênea definida sobre uma grade de 200×200 células. Células na cor verde representam a vegetação, células na cor vermelha estão queimando em $t = 100$ e a células pretas queimaram antes de $t = 100$. Como condição inicial a célula no centro da grade é levada à ignição em $t = 0$. Em (a) temos o padrão de queima dendrítico ($I = 0,25$). Percebemos em (b) $I = 0,30$ e (c) $I = 0,45$ que aumentos no valor de I produzem padrões mais sólidos.

3.1. Incluindo efeito de acúmulo de combustível

Como a probabilidade I associa-se à facilidade do fogo percolar a vegetação, podemos inferir diferentes valores de probabilidade de propagação para diferentes classes de acúmulo de combustível. Esses valores são aqui inferidos com base no valor do limiar de percolação e também usando o conceito que, quanto maior for o acúmulo de combustível na célula, mais intensamente ela queimará e mais provável será ela incendiar uma célula vizinha. A matriz resultante desses conceitos aparece na Figura 5(b) e contém as probabilidades de propagação entre cada par de classes de combustível. As colunas representam a classe de combustível da célula que está queimando, enquanto que as linhas correspondem às classes de combustível das células adjacentes que ainda não queimaram. Conforme observado nos incêndios mapeados, regiões queimadas há menos de um ano não possuem combustível seco acumulado suficiente para favorecer a propagação do fogo. Assim, o valor de probabilidade de propagação nessa classe de acúmulo é menor que o valor limiar I_c. Quanto maior o acúmulo de combustível, mais facilmente o fogo se propaga e, logo, maior será o valor de I. A Figura 5(a) mostra um exemplo de mapa de acúmulo de distribuição de combustível sobre a superfície. As classes de acúmulo aqui escolhidas são: queimado há menos de 1 ano (CA 0), queimado entre 1 e 2 anos (CA 1), queimado entre 2 e 3 anos (CA 2), queimado entre 3 e 4 anos (CA 3), e queimado há mais de 4 anos (CA 4).



(a)

INERTE \ QUEIMANDO	CA 0	CA 1	CA 2	CA 3	CA 4
CA 0	0.100	0.250	0.261	0.273	0.285
CA 1	0.113	0.253	0.264	0.276	0.288
CA 2	0.116	0.256	0.267	0.279	0.291
CA 3	0.119	0.259	0.270	0.282	0.294
CA 4	0.122	0.262	0.273	0.285	0.297

(b)

Figura 5. (a) Mapa de acúmulo de combustível para o mês de junho de 2002. (b) Probabilidade de propagação do fogo entre as classes de acúmulo de material combustível. As cores das classes corres correspondem às mesmas da legenda da Figura 5(a).

3.2. Incluindo efeito da velocidade e direção do vento

De forma geral, quanto maior é a velocidade do vento, maiores serão a intensidade da queima e a velocidade de propagação do incêndio na direção do vento. Ou seja, o vento induz uma tendência de propagação do fogo [16]. Conforme relatado em [10], podemos dividir os efeitos da velocidade do vento em três classes: ventos fracos (WS 0), com velocidades variando entre 0 a 5 km/h; ventos moderados (WS 1), com velocidades variando entre 5 a 20 km/h e; ventos fortes (WS 2), com velocidades maiores do que 20 km/h. Para cada uma dessas três classes de velocidades do vento, um *fator de tendência direcional* b_j é usado para modificar a propabilidade de propagação para as células vizinhas (ver Figura 6). Para WS 0 todos os fatores de tendência são iguais a 1. Valores maiores de velocidade do vento para WS 1 e WS 2 aumentam a probabilidade de propagação do fogo para células na direção do vento (ver Figura 6).

Os valores das tendências b_j são utilizados para modificar I e assim produzir uma *probabilidade direcional de propagação do fogo corrigida para o vento*, i_{wj}, na $j-$ésima direção ($j = 1, \ldots, 8$). O ajuste é uma probabilidade binomial cumulativa: $i_{wj} = 1 -$

Figura 6. Representação dos fatores de tendências direcionais usado para simular os efeitos da direção do vento. Sem vento ou vento fraco (WS 0: 0-5 km/h), vento moderado (WS 1: 5 - 20 km/h) e vento forte (WS 2: maior que 20 km/h). A flecha indica a direção do vento. Para WS 0 todos os fatores de tendências são iguais a 1 (sem tendência). O vento pode ser simulado vindo de qualquer uma das oito direções através da rotação da matriz de fatores de tendências até que o maior fator está posicionado a favor do vento. Figura adaptada de [10].

$(1 - I)^{b_j}$, onde b_j é o fator de tendência direcional apropriado, obtido através de uma matriz de oito possíveis direções relativas do vento (ver Figura 6). A posição da célula não queimada adjacente relativa a célula que está queimando determina qual fator de tendência é utilizado. Quando não existe vento, $b = 1$ e $i_{wj} = I$.

4. Simulações e Resultados

O modelo computacional foi desenvolvido no TerraME (*Terra Modelling Environment*), um ambiente de modelagem integrado a um sistema de informações geográficas, proposto por [5], que oferece uma linguagem de alto-nível para descrição de modelos baseados em espaços celulares.

Nesta seção simulamos dois cenários de incêndios ocorridos no Parque Nacional das Emas, conforme destacado na Figura 7 (indicados pelas setas vermelhas). Para se descobrir maiores informações sobre os incêndios, como data de início e fim e condições atmosféricas durante a ocorrência, combinaram-se informações dos mapas de cicatrizes de incêndios, dados de focos de calor gerados pelo INPE (Instituto Nacional de Pesquisas Espaciais), relatórios de ocorrência de incêndios emitidos pelo IBAMA (Instituto Brasileiro de Meio Ambiente e Recursos Renováveis) e dados meteorológicos registrados por duas plataformas de coleta de dados, uma no próprio Parque e outra na cidade de Chapadão do Céu, estado de Goiás, situada a cerca de 30 km da sede do Parque.

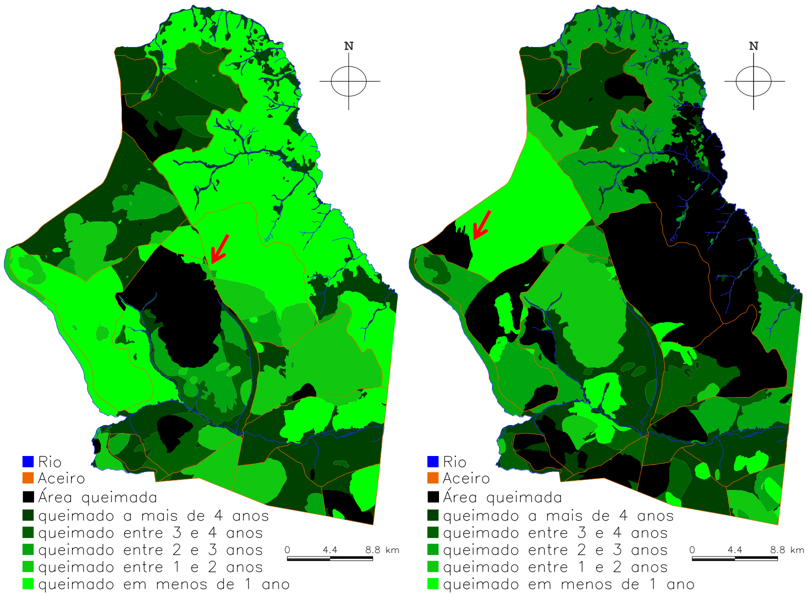
Através de consultas sobre o banco de dados geográficos obtiveram-se as seguintes informações dos incêndios:

A) Incêndio 1 (indicado na Figura 7(a)):

- Local de início: desconhecido;
- Data: 01-06-2000;
- Velocidade média do vento: 17 km/h (ventos moderados - WS1)
- Direção do vento: leste

B) Incêndio 2 (indicado na Figura 7(b)):

- Local de início: desconhecido
- Data: 09-07-2002
- Velocidade média do vento durante a ocorrência: 35 km/h (ventos fortes - WS2)



(a) (b)

Figura 7. Incêndios mapeados para o período (a) Junho de 2001 a Maio de 2002 e (b) Junho de 2002 a Maio de 2003 e suas respectivas classes de acúmulo de combustível para antes do período de mapeamento. Setas vermelhas indicam os incêndios simulados.

- Direção do vento: nordeste

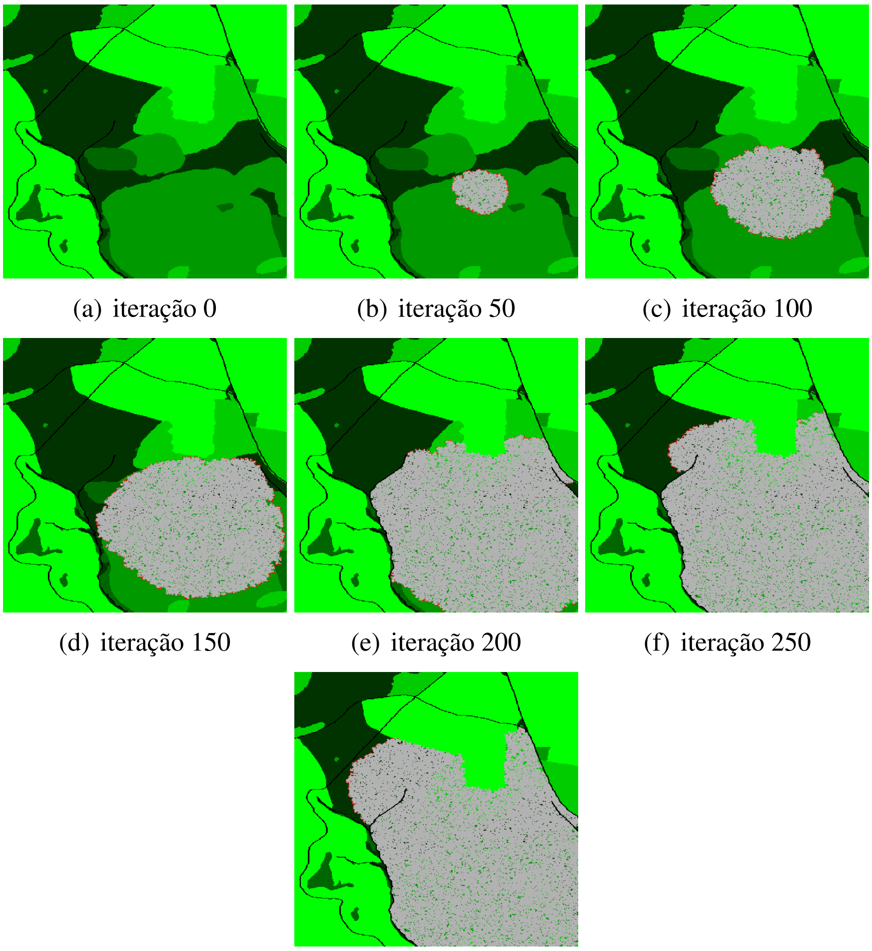
O local onde começou o incêndio é uma informação muito difícil de ser obtida, pois os incêndios são mapeados através de cicatrizes encontradas em imagens de satélite. Dados de focos de calor, obtidos de imagens termais de satélites de baixa resolução espacial, não são precisos o suficiente para se dizer onde ocorreu o foco.

A princípio, deve-se indicar como condição inicial qual célula (ou conjunto de células) está queimando. Essa escolha é feita aqui por hipótese, levando em conta a direção do vento e a extensão da área queimada. Após estabelecida a condição inicial, o modelo prevê o avanço da frente de fogo como um evento estocástico. As Figuras 8 e 9 apresentam a evolução dos cenários simulados para os dois incêndios.

Observa-se que o modelo apresenta distintos comportamentos da frente de fogo, dependendo da classe de distribuição de combustível na qual a célula se encontra. Células com alto acúmulo de material combustível apresentam maior facilidade de propagação do fogo, enquanto que células com baixo acúmulo favorecem a estagnação da propagação. As probabilidades direcionais induzem uma deformação da frente de fogo, favorecendo uma maior facilidade de propagação na direção do vento. Esse efeito pode ser evidenciado na Figura 9, onde a frente de fogo apresenta uma forma elíptica, indicando uma maior facilidade de propagação na direção do vento, que sopra para a direção leste.

5. Conclusões

O modelo utilizado neste trabalho se baseia na teoria da percolação. Os valores de probabilidade podem ser inferidos a partir de uma base de dados com histórico de incêndios ao longo dos anos e também com informações referentes aos fatores ambientais que,



(a) iteração 0 (b) iteração 50 (c) iteração 100

(d) iteração 150 (e) iteração 200 (f) iteração 250

(g) iteração 300

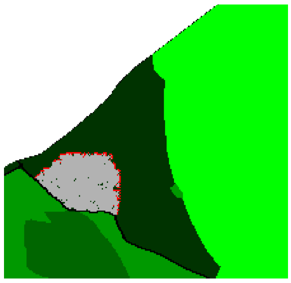
Figura 8. Sucessivos cenários de propagação previstos pelo modelo para o incêndio 1. Células cinza estão queimadas e vermelhas estão queimado no instante de tempo. As classes de distribuição de combustível são as mesmas conforme mostradas na Figura 7(a). As probabilidades de propagação utilizadas são mostradas na Figura 5(b).

evidentemente, influenciam na propagação do fogo. Os valores supostos para as probabilidades de propagação utilizados neste trabalho foram escolhidos também com base em evidências relacionadas às classes de anos sem queima. Regiões queimadas há menos de um ano não propagavam o fogo e, à medida que o tempo desde o último incêndio aumenta, a propagação é mais intensa. Esses valores serão substituídos por outros mais coerentes, obtidos através de correlações empíricas extraídas dos dados históricos e que levem em conta melhores efeitos do combustível, como curvas de acúmulo de biomassa seca, dominância de espécies e tipo de vegetação (ou tipo fitofisionômico).

A dominância, ou não, do capim-flecha (*tristachya Ieiostachya Ness*) é uma informação relevante para modelar a dinâmica do fogo no PNE. Esta gramínea, pelas suas características fenológicas e alta densidade, tem a capacidade de tornar o ambiente mais susceptível à ocorrência de incêndios intensos [17]. Assim, curvas de acúmulo de biomassa seca para regiões com e sem o capim-flecha fornecem informações adicionais para ajuste das probabilidades. Um outro fator que deve ser considerado é a classificação

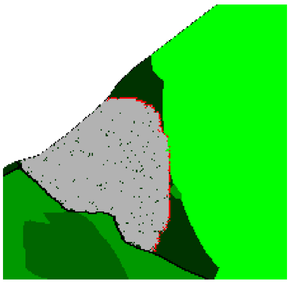
(a) iteração 0

(b) iteração 20

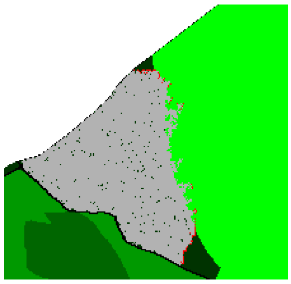


(c) iteração 40

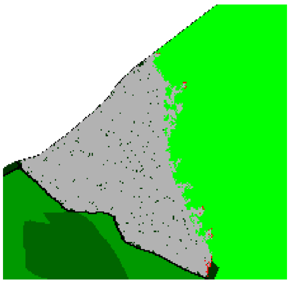
(d) iteração 60



(e) iteração 80



(f) iteração 100



(g) iteração 120

Figura 9. Sucessivos cenários de propagação previstos pelo modelo para o incêndio 2. Células cinza estão queimadas e vermelhas estão queimado no instante de tempo. As classes de distribuição de combustível são as mesmas conforme mostradas na Figura 7(b). As probabilidades de propagação utilizadas são mostradas na Figura 5(b).

da vegetação do Parque, como por exemplo, em função dos tipos fisionômicos.

Para efetuarmos simulações que possam prever a dinâmica do incêndio durante seu acontecimento faremos o mapeamento de cada iteração do modelo com o tempo real, incluiremos no modelo um *tempo de queima da célula* (que depende de fatores como velocidade do vento e classe de acúmulo de combustível da célula). A possibilidade do fogo avançar de uma célula para a sua vizinha só será avaliada após ultrapassado o tempo de queima da célula. Com esse caráter dinâmico, podemos incluir no modelo eventuais mudanças na velocidade e direção do vento, possibilitando a simulação de cenários mais realísticos que auxiliem na tomada de decisão em estratégias de combate à incêndios.

Percebe-se nas simulações apresentadas nas Figuras 8 e 9 que algumas células não queimaram durante a propagação do fogo. Isso se dá devido à natureza estocástica do modelo. A probabilidade atribuída se relaciona ao número de células vizinhas que serão ignizadas por uma célula que está queimando, a cada iteração. Para um máximo valor de probabilidade ($p = 1,0$), todas as células vizinhas serão levadas à ignição. Como a simulação é um processo estocástico, para chegarmos a uma conclusão sobre o cenário de incêndio, realizaremos um conjunto de simulações e contabilizaremos a freqüência de queima das células. O cenário será caracterizado por essa freqüência de queima, fornecendo uma informação acerca da probabilidade de uma determinada célula queimar dado um conjunto de condições ambientais.

A manutenção periódica da rede de aceiros é essencial para que os incêndios não atinjam grandes proporções do Parque, delimitando-os a blocos. Percebe-se no incêndio identificado pela seta azul na Figura 2(b) que o aceiro não foi eficaz. O estado da manutenção dos aceiros pode ser incluído no mapa de disponibilidade de combustível, associando-se a cada célula aceiro o tempo decorrido desde a última manutenção. Assim, podemos considerar ineficazes trechos de aceiro cujo tempo desde a última manutenção ultrapassou um ano.

O ambiente de modelagem TerraME mostrou-se bastante robusto para este tipo de aplicação, podendo então servir de base para o desenvolvimento de um aplicativo que auxilie na tomada de decisões em atividades relacionadas ao manejo do fogo em unidades de conservação.

AGRADECIMENTOS

Agradecimentos à Cordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo suporte à execução deste trabalho.

Referências

[1] BERJAK, S. G.; HEARNE, J. W. An improved cellular automaton model for simulating fire in a spatially heterogeneous Savanna system. **Ecological Modelling**, v. 148, p. 133–151, 2002.

[2] BURROUGH, P. A. **Principles of Geographical Information Systems for Land Resources Assessment**. Oxford: Clarendon Press, 1986.

[3] CÂMARA, G.; SOUZA, R.; FREITAS, U.; GARRIDO, J.; MITSUO, F. Spring: Integrating remote sensing and gis with object-oriented data modelling. **Computers and Graphics**, v. 15, n. 6, 1996.

[4] CÂMARA, G., R. SOUZA, et al. (2000). TerraLib: Technology in Support of GIS Innovation. **II Brazilian Symposium in Geoinformatics**, GeoInfo2000, São Paulo.

[5] CARNEIRO, T. G. d. S. **Nested-CA**: A foundation for multiscale modelling of land use and land cover change. 107 p. Tese (Doutorado) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2006.

[6] DUARTE, J.; TÉPHANY, H.; NAHMIAS, J. Percolation and fire spread. In: VIEGAS, D. X. (Ed.). **III International Conference on Forest Fire Research**. Luso, Portugal, 1998. v. 1, p. 361–362.

[7] FAVIER, C. Percolation model of fire dynamics. **Physics Letters A**, v. 330, p. 396–401, 2004.

[8] FRANÇA, H.; SETZER, A. **Regime das queimadas no Parque Nacional das Emas, GO: 1973-1995**. Relatório de Pesquisa FAPESP: Processo n. 95/2674-9, 1997.

[9] FRANÇA, H.; Ramos-Neto, M. B.; SETZER, A. O fogo no Parque Nacional das Emas. In: **Biodiversidade**. Brasilia, DF, Brasil: Ministério do Meio Ambiente, 2007.

[10] HARGROVE, W. W.; GARDNER, R. H.; TURNER, M. G.; ROMME, W. H. Simulating fire patterns in heterogeneous landscapes. **Ecological Modelling**, v. 135, p. 243–263, 2000.

[11] IERUMSALIMSCHY, T.; FIGUEIREDO, L. H.; CELES, W. Lua - an extensive extension language. **Software: Pratice & Experience**, v. 26, p. 635–652, 1996.

[12] KARAFYLLIDIS, I.; THANAILAKIS, A. A model for predicting fire spread using cellular automata. **Ecological Modelling**, v. 99, p. 87–97, 1997.

[13] MORAIS, M. **Comparing spatially explicit models of fire spread through chaparral fuels: A new model based upon the Rothermel fire spread equation**. Dissertação (Mestrado) — University of California, Santa Barbara, CA, USA, 2001.

[14] NAHMIAS, J.; TÉPHANY, H.; DUARTE, J.; LETACONNOUX, S. Fire spreading experiments on heterogeneous fuel beds. Applications of percolation theory. **Canadian Journal of Forest Research**, v. 30, p. 1318–1328, 2000.

[15] PASTOR, E.; ZÁRATE, L.; ARNALDOS, J. Mathematical models and calculation systems for the study of wildland fire behaviour. **Progress in Energy and Combustion Sciense**, v. 29, 2003.

[16] PYNE, J. S.; ANDREWS, P. L.; LAVEN, R. D. **Introduction to wildland fire**. 2nd. ed. New York, NY: John Wiley & Sons, 1996.

[17] Ramos-Neto, M. B.**O Parque Nacional das Emas (GO) e o fogo: Implicações para a conservação biológica**. 159 p. Tese (Doutorado) — Instituto de Biociências da Universidade de São Paulo, São Paulo, 2000.

[18] Stauffer D. and Aharony A., **Introduction to Percolation Theory** (Taylor and Francis, London) 1991.

Index of authors