HiperSIG: Hypermedia Technology Serving Geographic Information Systems

Celso Roberto Perez
Department of Informatics - UFPE
Caixa Postal 7851 - CEP 50732-970, Recife - PE - Brazil

and

Judith Kelner, Djamel Hadj Sadok, Ana Carolina Salgado
Department of Informatics - UFPE
Caixa Postal 7851 - CEP 50732-970, Recife - PE - Brazil

ABSTRACT

Geographical and Cartographic CAD/CAM applications, urban management and image processing applications, among others, represent a class of applications known as "spatial", which have been integrated into computational systems to form the so called "Geographical Information Systems" (GISs). An integration of hypermedia technology in such systems would take users from a world of static maps and manual presentations to an interactive environment. In this context, the present work analyzes the effect of using hypermedia technology for the spatial data manipulation, through navigation in the information hyperspace, as well as the retrieval and graphical presentation of such information. Hypermedia technology provides users with a set of tools them to access various models in order to explore separate scenarios. How this integration is achieved, its benefits as well as the role of each GIS component in the support for hypermedia technology are discussed in this paper.

Keywords: Hypermedia, Geographical Information Systems, Distributed Objects, Open Systems.

1. INTRODUCTION

Geographic Information Systems (GIS), as they are known, are software tools with support for the collection, management and analysis of high volumes of associated and spacely referenced data. They are used to help solve management problems, complex planning and research involving the manipulation of geographic data using geographic databases to store data (1). The stored data may be descriptive or graphical. Initially, GISs used to store both types of information in the form of internal simple files. With the rapid growth in the amount of spatial data being manipulated and the need for efficient access and retrieval of large amounts of data, this approach is currently being replaced by the increasing use of Data Base Management Systems (DBMSs).

These conventional DBMSs were not developed with geometric and spatial data manipulation in mind. Their embedded data and control structures, data definition and manipulation languages fail to properly handle or manipulate geographic information especially when considering aspects such as geometric and topological proximity (or neighborhood) characterizing spatial data as well as the inherent nature of geographic data. Some of the main characteristics of the latter include: complex relationships among information groups at different levels, the use of a non-conventional structure as a result of a different process of inheritance, the temporal aspect of the information, the use of specialized non-conventional information selection mechanisms, the high volume of data, etc. (2).

On the other hand, recent developments in the area of distributed processing seem to be moving towards new more flexible architectures based on distributed objects rather than the traditional rigid client server paradigm where all processing is concentrated at the servers. OMG's *CORBA* (3) and Microsoft's *OLE* (4) are two major players in the establishment of standards for distributed objects.

Considering current limitations of GISs in the storage and management of spatial data as well as new emerging multimedia distributed object processing systems, our work seeks to integrate both technologies giving GIS users transparent and efficient access and the ability to manipulate and present geographic data in their applications.

It is suggested that such integration is achieved through the use of *hypermedia technology*. This way, GIS users may navigate through, access and retrieve distributed data stored in distributed objects as shown in figure 1.

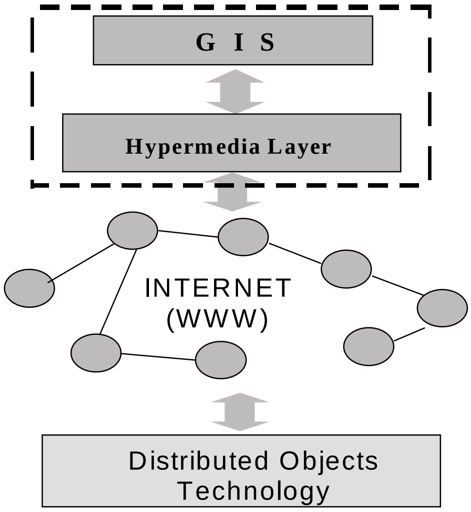

Figure 1: An integration scheme for GIS and distributed objects.

In particular, we concentrate in this work in dealing with one of the two aspects presented above, namely, how to enhance GIS systems using hypermedia functionality. This choice may be justified by a number of factors: hypermedia is an established technology for information access to distributed data, using navigation and search mechanisms. Hypermedia may be seen as data representation scheme that facilitates in the modeling of GIS data. Furthermore, it is an interactive technology based on objects.

2. HYPERMEDIA SYSTEMS

A hypermedia system may be seen as an information management technique where data is stored in a network of nodes. These may contain any type of digital information including text, images, audio, computer programs, video, etc.(5)(6). From the structural point of view, a hypermedia system is a direct graph with built in navigational mechanisms allowing users to move through it.

In a hypermedia system, the interface is an important component. This should be simple, with short response time and built to minimize the cognitive information overhead. In the context of GIS systems, hypermedia navigation through the search and browsing operations may be considered a serious limitation. Whereas a search operation may be performed when the user has identified the information sought, the browser is used to follow pre-established paths defined by the initial authoring process. Consequently, how easy the use of hypermedia application is going to be heavily depends on its author's ability to capture the semantics of the target application and to adequately organize the hypermedia graph structure. In addition to these two types of navigation mechanisms, document sequential reading is also (5).

Information retrieval in hypermedia systems, using associative links offers an alternative for data access from various points of a hyperdocument. This associative structure, not found in conventional methods, may also introduce some new problems of its own such as **difficulties in locating information** and **cognitive overhead** (5)(6).

3. GEOGRAPHIC INFORMATION SYSTEMS

These systems have, at different sophistication levels, the necessary functionality to manipulate and manage spatial data with includes data capture and entry, combining different data sources using manipulation algorithms generating new data types, data presentation and graphical and/or textual report generation.

A basic GIS characteristic is the support for spatial relations among geographical objects. The structure of these relationships reflects the topology information (neighborhood, proximity and pertinence) that may be established between geographical objects.

As earlier discussed, the main objective of a GIS is support for the storage, retrieval and analysis of geographic data. In order to achieve this, GISs define different functions, which may be divided into 5 levels: data collection, , pre-processing, data management, analysis and manipulation, and output generation (7). All of these mechanisms are implemented through two distinct components as part of a GIS. Figure 2 (8) shows these components as well as their relationship. Each system, depending on its objectives and necessities, may implement these components in different ways. However, all of these sub-systems are present within a GIS.

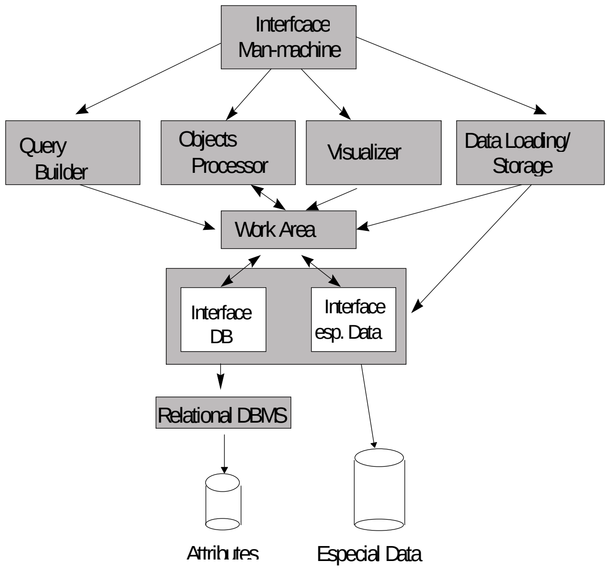

Figure 2: Anatomy of a GIS.

4. LIMITATIONS OF CURRENT GISs

Data sources required for integration into GISs are not only numerous but also very diverse in their origin and geographic location. Current database technology is not suitable for the efficient management of the complex GISs requirements. Considering the large volumes of data GISs are required to handle, coming from naturally distributed sources, distributed DBMSs may be seen as an initial solution towards the solution of data management. This, however introduces in our view other dimensions to the problem, such as: data communication, security, desktop technology, etc.

Other types of problems are related to the manipulation, processing and presentation of geographic data. Some of these include: **a-** relation and/or association problems among different types of information (it is also difficult to define 1:m relations between geographic objects), **b-** it is impossible to define "conceptual structures" between subsets of geographic objects, according to user requirements, **c-** inability to dynamically navigate geographical data sets from a particular application, or even to navigate through different levels of details of the same application **d-** cannot define work environments of sessions (including different data sets), which may stored and dynamically retrieved by different users, **e-** difficulty in defining the "procedures", "steps" or "guidelines" for the analysis and processing of certain data sets.

In other words, the major problems identified above by the users who have been interviewed may be divided into two groups. The first of these include all of those related to the manipulation of associated databases as well as tasks for the creation and loading of the cartographic bases. The second group include, on the other hand, those problems related to the manipulation, processing and presentation of data. According to user opinions, it is difficult as well as lengthy to get to learn how to manipulate any of these GISs as they require detailed knowledge in different areas such as the graphics, geometric computing, remote sensing, databases, etc. These limitations sometimes drive away a number of prospective users.

5. WHY WOULD WE INTEGRATE GISs AND HYPERMEDIA TECHNLOGY?

This integration facilitates the manipulation of both spatial and conventional data, through navigation mechanisms in the information hyperspace, as well as the retrieval and presentation of these multimedia information in appropriate output graphical devices.

GIS designers have, mostly until now, ignored hypermedia technology as a tool that may greatly improve their systems as shown in this work.

Incorporating hypermedia technology within these systems would provide new ways of seeing and managing application data, through users navigation among items of interest to them following pre-established links.

Unfortunately, many system designers still consider that hypermedia technology may only be used to access and manipulate documents. This "display oriented" behavior characterizes many of existing hypermedia systems. They have been designed to facilitate the authoring and creation of relations, information presentation and user navigation. The majority of system design methodologies were a specifically developed to optimize these services.

On the other hand, in computational applications, systems such as GISs, the document management is of secondary importance. Here, the hypermedia system should increase the activities and the processing of the data involved.

5.1. The role of each GIS Component

In our proposal, we consider GISs as having the following four components: a computational, a user interface, a hypermedia and finally a data management module. There are different ways for the implementations of these components, logical, such as the embedding the hypermedia functionality into the computational and user interface subsystems, developing a system as part of an application or separating the hypermedia module as a distinct component within a client server architecture for example. Implementation architectures are outside the scope of this work, but rather the logical behavior of these components. For clarity of our discussions, we assume for the rest of this paper that the hypermedia module is a separate one, and that the processing and hypermedia modules may gain access to the management module, which supports the objects storage, retrieval and management, as shown in figure 3.

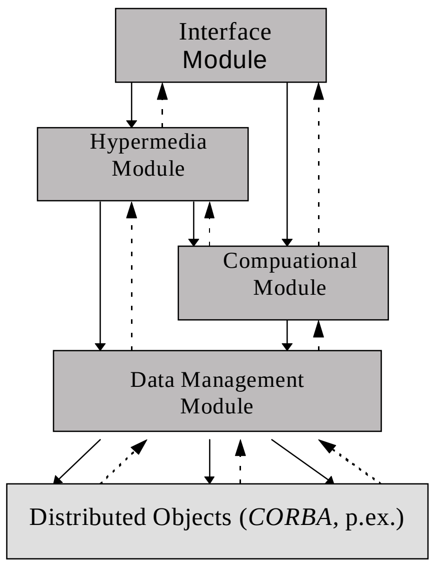

Figure 3: Module and their Relationships in the "*HiperSIG*" proposal.

Please note the difference between our "*HiperSIG*" proposed architecture and that of other GISs shown in figure 2. In "*HiperSIG*", associated data storage and management functions are performed within a single module, which turns transparent issues such as data types and their distribution. This is not the case in conventional DBMS used in GIS existing GIS architectures where there is a division where there is a division between the data supported by conventional DBMSs and spatial data. Another important difference is that the functionality which in figure 2 appear separated, are now integrated in the processing (computational) module.

The hypermedia module manages all hypertext functionality on the behalf of the application. Such functions (navigation, annotation are structure and visual oriented) will be built using basic blocks of nodes, links and anchors. This way, the hypermedia module should internally represent the application objects in terms of these constructs: nodes: maps, models variables, data sets, CAD objects, computed results, etc., *links*: represent relationships, for example, between data sets, between a calculation and an associated annotation or between components of a composite map or CAD objects, anchors: delimit the start and end of links. Depending on the type of the link, anchors may refer to an entire object or a specific object within the contents of a node, such as a segment of a coast in a complex map.

Managing hypermedia functions in a GIS application requires cooperation between the interface and computing (processing) modules. First the hypermedia module should communicate with other modules. This implies that there should be paths and protocols to support the flow and the interpretation of messages. Furthermore, the interface module should present multimedia information, allowing users to select parts of this information by exchanging the appropriate messages with the hypermedia module as a result of some event.

The hypermedia module would intercept messages as they come from the computational module containing presentation requirements. It should therefore know the format of these application messages. Through arguments, key word searches, lexical analysis or the content, unique identification of objects, etc., the hypermedia module identifies application object references and maps the "*stand alone*" objects to node. When a user selects an anchor, the interface passes its identifier to the hypermedia module. This determines the link associated to the anchor, asking the computing module for adequate processing or the data management module to retrieve the associated object. The database query or that of the application scheme structure, returns the object sought and determines if this has any links. If this was the case, the hypermedia module maps the object references to anchors.

5.2. Hypermedia Systems Functionality

These include navigational characteristics, of annotations, structure and visual oriented. In order to support this, objects manipulated by the GIS must be defined which are considered as nodes as well as their links. Next, we examine and evaluate some proposals and justify the use of hypermedia to overcome the limitations of GISs.

5.2.1. What de we link? The philosophy of maximum Access

We start by considering what should be linked in a GIS application. What should the hypermedia represent as anchors and links in this computational environment.?

SO far we saw hypermedia as a "maximum access" technology giving the user the freedom to access and explore information and meta-information. Any item of interest to the user is a suitable candidate to be an anchor. In the case of GISs, this includes: **1-** objects traditionally linked in applications non computational, **2-** objects for which meta-information of the type: definition, attribute values and even executable programs, commands and parameters are available, **3-** results from any computational operations (calculus, clarifications, composition of individual objects etc.).

On the other hand, may need to access various associations or links. Possible links include: **1-** connections between an object and its meta-information (a map and its scale information, date of creation, author, tools used for its elaboration, etc.), **2-** operations connecting input and output, **3-** structural and group relationships (components of a geographical phenomenon, such as complex satellite image), **4-** different visions of the same object (video and network sewage, for example), and **5-** spatial or temporal relationships (the sewage network in times when, for example, the tide is high or low).

5.2.2. Automatic Linking through Structural Relations

Manual linking em dynamic computational systems, such as in GISs, is partly not feasible due to the large number of application elements and since many of these elements do not exist when the system designer originally build his/her application. This is the case of computational results for example.

When the computational module sends a message to the user interface module, the hypermedia module should automatically analyze this message and dynamically map the instances of " and "structural links" to application objects. This eliminates all manual links, except to those objects and relationships that cannot be generalized (non structural links will be discussed in the annotations functionality).

Let's examine the example shown in figure 4, when a geographical region is presented, the user may gain access to some or all of its structural links, such as the geographical characteristics (rivers, roads, etc.), the set of spatial data related to the management of water, energy distribution, etc., the evaluation of the impact of natural phenomena, etc. These structural relationships, which in the HDM model are as authoring in the large (9), should be equal in the application scheme. System and project designers may declare these relationships in different forms, including rules of specialized systems, or even as prolog predicates.

5.2.3. Dynamic Regeneration

Within environments that manipulate geographic information, many of the discussed hypermedia characteristics face the challenge of dynamic regeneration. For instance, a GIS user wants to know if there is correlation between the family income and the amount of money the city spends on sewage system in a given region.

The user selects a region of interest to him/her in the displayed map and invokes the appropriate processing, by choosing the items and variables of interest from the menus or directly from the display. The user then may annotate his/her reasons for doing this calculation. Later on, while preparing a report for example, Later on, while preparing a report for example, the user may request the results as well the associated annotations, without the need to look for this information in the original map nor to specify again the parameters, etc. In other words, the hypermedia module would a call mechanism that would enable it to automatically regenerate the calculation results. The hypermedia module will have to deal with the challenge of reallocate annotations previously declared in regenerated documents, specially when content changes between versions.

5.2.4. Navigation Functionality

As discussed in section 2, navigation allows users to move freely between locations in a hypermedia network of (analysis, maps, data sets, etc.). Characteristics to navigation include *browsing*, *backtracking*, querying based in components and structures.

a. ***Browsing:*** allows a user to select and follow a link. In hypermedia GISs, following an operational link may cause the computational module to generate (or refresh) the final point (the target). Browsing updates the system log, used by the hypermedia machine for *backtracking*. Entries in the log include the nodes of origin, links and anchors, and context information. The entry values may contain hypermedia and computational module identifiers, parameter values for procedure calls institution, and any other information necessary to regenerate the output point when backtracking is activated.

b. ***Backtracking:*** allows a user to trace back his/her steps. It uses the session log to enable the return to previously visited nodes, but in its current state. For example the analysis of rain fall discussed in the section of dynamic regeneration. This is a sub-task within a bigger project which requires knowing separate rain fall estimates of for each of the regions. The user may start with a high level view of the regional hydraulic management, analyzing spatial data at a scale of 1:24000. The user may start looking for more details related to the rain drainage system, each time changing, the presentation scale. Each step a user performs calculations and annotations and may obtain results. Although this scenario seems similar to navigation based on hypertext documents, it is however different in that navigation is a direct result of the of the computations performed dynamically. These resulting nodes are computed dynamically, where the content is a factor of the current context, presenting a selected link. *Backtracking* allows the user to see the calculations, at each level of the scale, before continuing with the next task.

Queries based on Components and Structures: these types of queries may take the user to either a point for *browsing* information space, or to a subset of application data. The queries traditionally supported by computational module may be enhanced with hypermedia functions (for example, allowing users to select the elements that specify the query, or even to define a "*guide tour*" before establishing how it is going to be completed). In addition, user may query the hypermedia met-information base, with structural queries, combining therefore application components with the hypermedia structure.

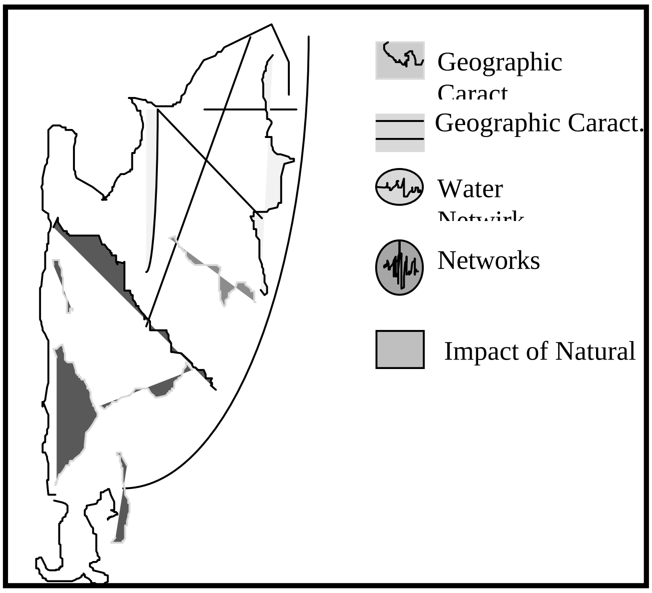

Figure 4: Example of structural links associated to a geographic region

5.2.5. Annotations

Annotations offer users a way to add information to the GIS application hypermedia structure in addition to the one inferred from the hypermedia module using application structure. Users and designers may declare non structural links in the form of annotations representing links that may not be inferred as well as ad-hoc ones. Such comments may remind us of annotations and personal analysis, offering meta-information to the author or other users.

Examples of meta-information subject to annotations are: object classes (a general comment about the regression analysis or one that is specific to any data value used in this analysis), or the hypermedia (comments related to a computational link).

Users may also associate multiple analysis using the same parameter values for each one of them, such as when establishing a link between the rain fall analysis and coastal corrosion at the same maritime region. These types of non structural relations may connect these objects only if there are no predefined structural relations between them.

Furthermore application designers will have to decide whether to distinguish between annotating specific object instances and (commenting a data value within a specific calculation, for examples) or to annotate a general object (continuing similarly to the previous example, the comment should be associated with any presentation of the data value in any, existing or generated, document, map, data set, etc.). In order to support such non structural links, when a message has been converted from the computational to the user interface module, the hypermedia module should find all possible annotations and structural nodes and anchors. The search of these components, which may not be identified in another way, may also be triggered within regenerated documents.

If the document did not substantially change, between different generations, therefore the anchor should remain near the same position or should have the same display value. Another possibility, may be that the project designers may restrict users in that they can only annotate objects that have been declared as structural, i.e. those that the hypermedia module may easily locate.

5.2.6. Vision Oriented Functionality

Overviews, paths and guided tours *offer* interesting alternative ways of seeing and navigating through an application.

a. Overviews: display application components such as anchors with lines interconnecting them representing their relationship. Furthermore, an *overview* may take the shape of a coherent display such as a map. When dealing with computational applications, overviews offer access to computations, as well as static nodes. Users, not only, may the global context, but may also invoke a computation or analysis when selecting a given anchor.

A major limitation of *overviews* however, is that they generally cannot reflect correctly all inter-data and inter-document structural relations, as well as all possible computations. For instance, spatial data sets may contain hundreds more relations, much if which only exist within given views and criteria. As pointed out earlier, this remains a major problem yet to overcome when manipulating spatial data in general and specifically in the case of GISs.

A way to solve this problem, would be the use of hypermedia technology to build abstractions for relations presented in the *overviews*. More importantly, such abstractions may be presented as display *links* without actually hiding its content. Doing this automatically requires the spatial analysis of overview contents. This is illustrated in the example shown in figure 5, the *HiperSIG* presents a map of a specific region in Brazil, where medium rain fall rates appear in different colors. In this same figure, are also shown graphic bars showing the average rate for of water drainage for each of these regions. The latter have been generated by the computational module and are used by the hypermedia module as link marks representing information from respective analysis.

Each one of these links may provide access to information such as: the original data used in the analysis, a description of collection methods used to get the data used in this analysis, the frequency and collection time interval, and even a detailed analysis of the draining process such as by street or area, etc.

b- Paths and "guided-tours":. Offers users a recommended way to navigate through in the hypermedia network. The paths allow users to freely leave and rejoin them. "Guided-*tours"* on the other hand restrict users to the anchors links and annotations of the path. These paths may guided novice users, just like a *"hands-on"* demo or as a training tool for the execution of analysis. The combination of this with annotations of all the steps executed within an analysis results in the documentation of the whole process of analysis and the justifications for each of the decisions taken. This hypermedia systems feature is a natural and suitable solution for GIS manipulation and usage problems by users.

5.3. Proposal benefits and the required coordination for complete integration

Table 1 lists, in concise way, the limitations found in current GISs earlier identified in section 4. It also shows the hypermedia features we provided and discussed as solutions to these problems. On the other hand, in order to implement GISs with hypermedia technology, there would be no need to rebuild all hypermedia functionality from scratch. Hypermedia models such as DEXTER (10) are in our view a good starting point, through *frameworks*, data structures and procedural guidelines for the building of nodes, links and anchors, as well as the support for navigation. Different systems and architectures may offer a suitable base for GIS applications. In the general case of common computational applications, we may refer to works from Boursier and Dbouk (11)(12), we may refer to works from Boursier and Dbouk (11)(12), as a computational example in the hypertext area, and to work from Bieber et al. (13) as a hypertext example in the computational area.

6. CONCLUSIONS

This work suggests the integration of GISs and hypermedia technology to produce a "*HiperSIG*". There is a clear lack of research in this direction and a number of issues raised in this document need to be addressed. A major problem discussed here, was the relatively large volume of information GISs usually deal with. Linking object documents becomes a problematic task.

Consequently, traditional methods oriented towards the identification display and link establishment as used in hypertext systems are in fact not feasible. Adaptive and dynamic techniques based on structural acknowledgment may be used to enhance computational applications with navigational characteristics, annotations, search operations that both structure and view oriented, etc.

With the application of hypermedia functionality to the computational domain, as in the case of GIS, according to the "maximum access" philosophy , these would assist users in the development of applications that are easy to understand and highly reliable and confidentiality. In order to achieve this, there remains considerable work to be undertaken. Some of this work may: **a-** the study of new dynamic techniques for the establishment of links, based in the geographic applications structural acknowledgment, **b-** further research into non structured links (annotations) using non conventional media such as video and audio, **c-** to evaluate the possibility and effectiveness of using hypermedia technology in order to solve GIS problems especially when dealing with distributed data and processing platforms, **c-** to deal with synchronization (presentation) issues resulting from the use of hypermedia systems combined with distributed GIS systems, **d-** to consider the advantages of using interpreted languages such as Java and *scripts*, *(CGI (Common Gateway Interface)*, and even high level semantically rich APIs *(Application Programming Interface)*, and integrate all of these ingredients into the implementation of *GIS + Hypermedia + Distributed Objets.*

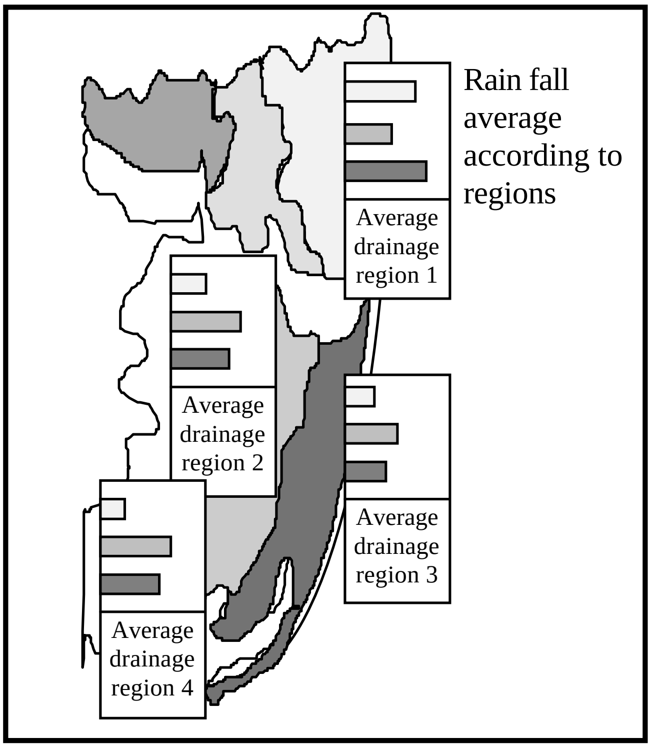

Figure 5: Abstraction of Link "drainage" per average rain fall.

	LIMITATIONS OF CURRENT GISs	SOLUTIONS USING HYPERMEDIA
IN RELATION TO DATA BASES	- Need to share geographic data (lack of a common database, of communication standards and/or interchange)	- The use of multimedia distributed databases, with hypermedia standards for the document interchange including *HTML, SGML, ODA/ODIF, HyTime, MHEG, etc.*
	- Inefficient, slow and repetitive database table creation and loading processes	- Reuse of data available through a network
	- Difficulty in the manipulation of stored data requiring specific database knowledge.	- Abstraction, user transparency through the use and definition of the primitives nodes, links and anchors.
IN RELATION MANIPULATION PROCESSING AND DATA PRESENTATION	- Limitations in linking/associating different information types	- Nodes definition, different types of links (structural, "move-to", *zoom*, views, indexed, executable, etc.) and anchors, with the use of the "maximum access" philosophy.
	- It is impossible to define conceptual structures between geographic data sets, accurately relation within a particular application.	- Definition of non structural links (annotations) with the possibility of automatic links through structural r elations.
	- No support for dynamic navigation through geographic data sets of a particular application.	- the use of navigation mechanisms such as *browsing*, *backtracking* and component and structure based search; and view oriented functionality such as *overviews.*
	- Does not define work environments or sessions which may be saved for dynamic retrieval by users.	- Dynamic regeneration of processing and/or work sessions, non structural links
	- Difficulty in defining procedures steps, guidelines or procedures for the analysis and processing of certain data types.	- View oriented functionality (*overviews*, paths and *guided-tours*)

Table 1: A summary of current GISs problems and their solutions using Hypermedia

7. REFERENCES

(1) United State Environmental Protection Agency GIS Program. http://www.epa.gov/ngispr/gis.html

(2) Haas, L. and Cody, W. "Exploiting Extensible DBMS in integrated Geographic Information Systems". In Proc. 2nd. Symposium Spatial Database Systems, pp. 423-449. Springer Verlag Lecture Notes in Computer Science 525. 1991.

(3) CORBA, http://www.omg.org.

(4) *OLE: Object Linking and Embedding*, http://www.microsoft.com.

(5) Rada, Roy "Hypertext: From text to expertext". Mc Graw-Hill, 1991.

(6) Nielsen, J. "Hypertext and hypermedia", Academic Press, Inc. 1990.

(7) Star, Jeffrey, et.al. "Geographic Information Systems, an introduction", Prentice Hall, Englewood Cliffs, 1990.

(8) Câmara, Gilberto "Anatomia de um GIS". Em Fator GIS, 1994, No 4, pp. 11-15.

(9) Schwabe, Daniel, et.al. "Introdução aos Sistemas e à autoria hipermídia". Caderno de textos. In Escola Brasileira Argentina de Informática, 6 (EBAI), jul 1993, Córdoba, Argentina.

(10) Halasz, F. Schwartz, M. "The Dexter Hypertext reference model". Communication of ACM. Vol. 37. No 2. Feb. 1994, pp. 30-39.

(11) Boursier, P. et.al. *"Hypermedia techniques and spatial relationships in HyperGeo Dynamic Map Applications"*. Em Joint European Conference on Geographical Information Systems, Barcelona, Spain, march 1996.

(12) Dbouk, M. et.al. *"A multi-representation and multi-scale data model for hypermedia GIS applications"*. Em Proceedings of Geoinformatics´96, USA, April 1996, pp. 41-50.

(13) Bieber, M. "Automating hypermedia for decision support". In Hypermedia vol. 4, no. 2, 1992, pp. 83-110.