Geohash Index Based Spatial Data Model for Corporate
基于 Geohash 索引的企业空间数据模型

Conference Paper · August 2015
会议文件 2015 年 8 月

DOI: 10.1109/ICEEI.2015.7352548
Doi: 10.1109 / ICEEI. 2015.7352548

5 authors, including:
5 位作者，包括:

Iping Supriana
伊平苏普里亚纳多迪法
Bandung Institute of Technology
万隆理工学院万隆理工学院

288 PUBLICATIONS **430** CITATIONS
288 出版物 430 引文 10 出版物 15 引文

SEE PROFILE
参见配置文件

Dody Dharma
Bandung Institute of Technology

10 PUBLICATIONS **15** CITATIONS

SEE PROFILE

Dicky Prima Satya
迪奇·普里马萨蒂亚
Bandung Institute of Technology
万隆理工学院

5 PUBLICATIONS **14** CITATIONS
5 出版物 14 引文

SEE PROFILE
参见配置文件

Some of the authors of this publication are also working on these related projects:
这本书的一些作者也在从事这些相关的项目:

Geospatial Search Engine View project
地理空间搜索引擎视图项目

FACE RECOGNITION GENERIC TO SPECIFIC FEATURE REPRESENTATION AND RECOGNITION STRATEGY View project
人脸识别通用特征表示和识别策略视图项目

The 5th International Conference on Electrical Engineering and Informatics
2015 August 10-11, 2015, Bali, Indonesia
2015 年 8 月 10 日至 11 日，印度尼西亚巴厘岛，第五届国际电子工程与信息学会议

Geohash Index Based
基于 Geohash 索引的数据挖掘方法
Spatial Data Model for Corporate
企业空间数据模型

Iping Supriana Suwardi, Dody Dharma, Dicky Prima Satya, Dessi Puji Lestari
伊平苏普里亚纳苏瓦迪，多迪达尔马，迪奇普里马萨提亚，德西普吉莱斯塔里
School of Electrical Engineering & Informatics, ITB,
电子工程与信息学院,
Bandung, Indonesia
万隆
{iping, dody, dicky, dessipuji}@informatika.org
{ iping，dody，dicky，dessipuji }@informatika. org

Abstract— **Spatial data wrapped and processed by an application known as Geographic Information Systems (Geographical Information System / GIS). In desktop based GIS, the spatial information services only occurs when a variety of basic data has been loaded into the applications database. The development of information technology has provide ready to access, worldwide scale, web based mapping system, like Google Maps. The addition of the information content to the map can also be done easily by any user into the system. Until now, both the basic data provided by Google Maps or data that is added by users are loosely connected, meaning that there are minimum linkage between data. Thus, the data for greater benefit of a corporation or government where the data is closely related to each other still yet to be served. As spatial data being managed are voluminous, the scalability of querying performance will be a challenge. To anticipate this, we describe an improvement that built on top of our proposed spatial data model. We used a special data which derived by interleaving bits obtained from latitude-longitude pairs of a spatial data, the string called geohash. A geohash can be used as an index of every object in Spatial data table. The longer the Geohash string, the more precise the bounding box around the location it references. This approach will improve the performance of querying process of a single or even collection of spatial data in the data table of corporate GIS. The main study of this research is to provide information services along with the availability of a variety of basic spatial data owned by Google Maps. This paper highlights our recent effort in theoretical and applied research in spatial data management.**

抽象—由地理信息系统(geographicinformationsystem/GIS)应用程序包装和处理的空间数据。在基于桌面的 GIS 中，空间信息服务只有在将各种基本数据加载到应用程序数据库中时才会发生。信息技术的发展已经提供了全球范围内便捷的基于网络的地图系统，就像谷歌地图一样。将信息内容添加到地图也可以由任何用户轻松地进入系统。到目前为止，谷歌地图提供的基本数据或用户添加的数据都是松散连接的，这意味着数据之间的连接极少。因此，为了公司或政府的更大利益而提供的数据，在这些数据彼此密切相关的情况下，仍然有待提供。由于管理的空间数据量庞大，查询性能的可扩展性将是一个挑战。为了预见到这一点，我们描述了在我们提出的空间数据模型之上进行的改进。我们使用了一种特殊的数据，这种数据是通过从空间数据的经纬度对获得的交错位得到的，这个字符串称为 geohash。Geohash 可以用作 Spatial 数据表中每个对象的索引。Geohash 字符串越长，它引用的位置周围的边界框就越精确。这种方法将提高企业 GIS 数据表中单个甚至集合空间数据的查询过程的性能。这项研究的主要目的是提供信息服务以及谷歌地图所拥有的各种基本空间数据的可用性。本文着重介绍了我们在空间数据管理理论和应用研究方面的最新成果。

Keywords—geohash, geographical information system, spatial data model, mapping system

关键词—geohash、地理信息系统、空间数据模型、制图系统

I. INTRODUCTION
引言

Modern technology now permits improved acquisition, distribution, and utilization of geographic or geospatial data (Craglia, 2006). There are many popular web based mapping technology. By using the technology, people are able to get any information based on earth coordinates. The technology are very sufficient for personal and daily use, but there is a need to utilize it for corporate or governmental data needs.

现代技术现在允许改进地理或地理空间数据的获取、分配和利用(Craglia，2006)。有许多流行的基于 web 的地图技术。通过这项技术，人们可以获得任何基于地球坐标的信息。这项技术对于个人和日常使用是非常充分的，但是需要利用它来满足公司或政府的数据需求。

Geospatial information is critical to promote economic development and improve stewardship of natural resources. The use the mapping technology for corporate or governmental need will involve the preparation of the data and the relationship between geographic data in a very detailed level (e.g. to support the interests of plantation management, taxation, city layout, quality of roads, quality of rivers, etc.).

地理空间信息对于促进经济发展和改善对自然资源的管理至关重要。为了公司或政府的需要而使用制图技术，将涉及到非常详细地编制数据和地理数据之间的关系(例如，支持种植园管理、税收、城市布局、道路质量、河流质量等方面的利益)。

Different corporate will have a different data needs, and different data representation.
不同的公司会有不同的数据需求，以及不同的数据表示。

The development of information technology has provide ready to access, worldwide scale, web based mapping system, like Google Maps. The addition of the information content to the map can also be done easily by any user into the system. Information added by these users may also be registered into Google Maps as information that can be read by any other users. So that a variety of important information that is needed by the public can be easily searched and known, for example the location of hospitals around the world, the location of government institutions, sports location or any location by contributing the data.

信息技术的发展为人们提供了便捷的、全球范围的、基于网络的地图系统，比如谷歌地图。将信息内容添加到地图也可以被任何用户轻松地添加到系统中。这些用户添加的信息也可以作为其他用户可以阅读的信息注册到谷歌地图中。因此，公众需要的各种重要信息可以很容易地被搜索和了解，例如世界各地医院的位置、政府机构的位置、体育场所或任何提供数据的地点。

Voluminous geographic data have been, and continue to be, collected with modern tools or gadget like GPS or smartphone, location aware services, and crowdsourcing internet geospatial information. The scale of geospatial data is often too large and too unstructured with not enough relationship explanation.

大量的地理数据已经并将继续通过现代工具或小工具收集，如全球定位系统或智能手机、位置感知服务和众包互联网地理空间信息。地理空间数据的规模往往过大，结构过于松散，没有足够的关系解释。

Until now, both the basic data provided by Google Maps and data that is added by users are loosely connected, meaning that there are minimum linkage between data. Thus, the data for greater benefit of a corporation or government where the data is closely related to each other still yet to be served. For example, State Electricity Company (PLN) needs to manage spatial data like electricity substation location, the location of electricity poles, wires between poles and the location of the customer. In other words, web based mapping system like Google Maps can not be utilized effectively for the sake of analysis, especially for the benefit of corporate analysis. In the example above, when the electricity substation hit by a lightning strike, analysis is required to identify any customer who will be affected by a power outage; and then the data of affected customers will appear as a certain color in the Google Maps display

到目前为止，谷歌地图提供的基本数据和用户添加的数据都是松散连接的，这意味着数据之间的连接是最小的。因此，为公司或政府带来更大利益的数据，如果这些数据彼此密切相关，那么这些数据仍然有待提供。例如，国家电力公司(PLN)需要管理空间数据，比如变电站的位置、电线杆的位置、电线杆之间的线路以及用户的位置。换句话说，像谷歌地图这样的基于网络的地图系统不能有效地用于分析，特别是对于企业分析的好处。在上面的例子中，当变电站遭到雷击时，需要进行分析以确定哪些客户将受到断电影响;然后受影响客户的数据将在谷歌地图显示中以某种颜色显示

In order to provide facility for services model above, this research proposed. The main study of this research is to provide information services along with the availability of a variety of basic information owned by Google Maps. This paper highlights our recent effort in theoretical and applied research in spatial data management. We are proposing a new model of spatial data management that will accept spatial data structure and create relationship to basic spatial data.

为了给上述服务模型提供便利，本研究提出了。这项研究的主要目的是提供信息服务，同时提供谷歌地图所拥有的各种基本信息。本文着重介绍了我们在空间数据管理理论和应用研究方面的最新成果。我们提出了一种新的空间数据管理模型，它接受空间数据结构，并与基础空间数据建立关系。

II. STATE OF THE ART
最新技术

The development of information technology enables organizations to be able to perform digital geographic data
信息技术的发展使各组织能够执行数字地理数据

management. Digital geographic data has been widely used to help solve critical problems through geographic analysis in various industrial sectors. The use of geographic data, among others, can help the analysis of geographic data for marketing optimization through segmentation of customer data; assist urban management and transport; management of natural resources; simulation; and environmental conservation [1]. The use of geographic data today has penetrated into various mobile devices to assist a person in getting the services closest personal information

管理。数字地理数据已被广泛用于帮助解决关键问题，通过地理分析在各种工业部门。地理数据的使用，除其他外，可以帮助分析地理数据，通过对客户数据进行细分，优化销售;协助城市管理和交通;管理自然资源;模拟;环境保护[1]。如今，地理数据的使用已经渗透到各种移动设备中，帮助人们获得最接近个人信息的服务

In general, geographic data wrapped and processed by an application known as Geographic Information Systems (Geographical Information System / GIS). GIS has the ability to perform spatial analysis, which is a set of analysis methods that require access to the attributes and location information of an object (Goodchild (1988) in [2]). Spatial data consists of graphic maps correlated with the attribute table, which enables users to visualize and perform queries quickly. In implementation, a number of GIS has a different way of reading and storing spatial data.

一般来说，地理数据由称为地理信息系统(GeographicInformationSystem/GIS)的应用程序进行打包和处理。Gis 具有执行空间分析的能力，空间分析是一组需要访问对象的属性和位置信息的分析方法(Goodchild(1988)，见[2])。空间数据由与属性表相关的图形地图组成，这使用户能够快速地可视化和执行查询。在实现过程中，许多 GIS 具有不同的读取和存储空间数据的方式。

In the early stages, graphical information service is created as computer aided design applications, which provide the service of making a line on the computer to design different drawing interest. Developments of graphical information service is continue specific for mapping which provide the service of making a line on a computer for the purpose of making a map that runs on the local computer (desktop). Further development is providing services in creating line and curvature more easily again that not just runs on desktop platforms, but also runs on a web platform.

早期的图形信息服务是作为计算机辅助设计应用而创建的，它提供在计算机上绘制线条以设计不同的绘图兴趣的服务。图形信息服务的发展继续专门用于制图，这种服务提供在计算机上绘制线路的服务，目的是制作在本地计算机(桌面)上运行的地图。进一步的发展是提供服务，创建线和曲率更容易再次不仅运行在桌面平台，而且运行在一个网络平台。

Openshaw (1987) commented that such systems are basically concerned with describing the Earth’s surface rather than analyzing it (Openshaw(1987) in [3]). In such systems, the spatial information services only occurs when a variety of basic data has been loaded into the applications database. It is a real difficulty for a corporate to build such system, because they have to supply the basic spatial data, their specific spatial data, and define the relation.

奥本肖(1987)评论说，这样的系统基本上是描述地球表面，而不是分析它(奥本肖(1987)在[3])。在这样的系统中，空间信息服务只有在将各种基本数据加载到应用程序数据库中时才会发生。由于企业需要提供基本的空间数据、具体的空间数据，并且需要定义空间数据之间的关系，因此建立这样的系统是企业面临的一个实际难题。

In the Google Maps service, information can be directly accessed by the user without worrying the entry of basic spatial data. The basic data likes coordinates, roads, borders, and so on is already provided. In addition, Google Maps has facilitate user that wants to enter the data for the user's own interests. System development opportunities also become possible, because Google Maps has the basic functions that can be accessed by other applications (via the Google Maps API facility). So that developers can develop applications which is an extension of the use of a variety of basic information that Google Maps for corporate needs.

在谷歌地图服务中，用户可以直接访问信息，而不必担心基本空间数据的输入。已经提供了坐标、道路、边界等基本数据。此外，谷歌地图还为用户输入数据提供了便利。系统开发机会也成为可能，因为谷歌地图具有其他应用程序可以访问的基本功能(通过谷歌地图 API 功能)。因此，开发人员可以开发应用程序，这是谷歌地图用于企业需求的各种基本信息的扩展。

There seems to be widespread agreement in the GIS community on two simple propositions: that as technology, GIS has the potential to support many different type of analysis and that this potential has not yet been realized [5].

地理信息系统界似乎普遍同意两个简单的主张:作为技术，地理信息系统有可能支持许多不同类型的分析，而这种可能性尚未实现[5]。

We now can obtain much more diverse, dynamic, and detailed data. Generally speaking, geography and spatial sciences have moved from a data-poor era to a data-rich era. The availability of vast and high resolution spatial and

我们现在可以获得更加多样化、动态和详细的数据。总的来说，地理学和空间科学已经从数据匮乏的时代进入了数据丰富的时代。广阔和高分辨率的空间和

spatiotemporal data provides opportunities for gaining new knowledge and better understanding of complex geographic phenomena, such as human-environment interaction and social economic dynamics, and address urgent real world problems, such as global climate changes and pandemic flu spread [4].
时空数据提供了获得新知识和更好地理解复杂地理现象的机会，如人类与环境的相互作用和社会经济动态，并解决紧迫的现实世界问题，如全球气候变化和大流行性流感的传播[4]。

To gain those knowledge, there is a need to analyze the spatial data. But in order to do that, the data must be presented in a structure with clear relations between data from different geographic information infrastructure.
为了获得这些知识，需要对空间数据进行分析。但是为了做到这一点，数据必须以一种结构的形式呈现，这种结构必须在来自不同地理信息基础设施的数据之间有明确的关系。

To increase efficiency and interoperability of geographic information infrastructure, many regional and global initiatives work in the establishment of open standards and agreements. Content is managed by means of regulated and standardized service types. This imposes a distinct life cycle of geospatial content in distributed environments which can be be described in four steps illustrated in Figure 1. First, content must be made available in distributed system, i.e., content must be published in standard services like discovery and access services. Second, users need to discover content which will be finally accessed by using these services (third step). Finally, users process the content and generate new content, which should be integrated and published in the distributed system closing cycle [5].
为了提高地理信息基础设施的效率和互操作性，许多区域和全球倡议致力于制定开放标准和协议。内容是通过规范和标准化的服务类型来管理的。这为分布式环境中的地理空间内容提供了一个独特的生命周期，如图 1 所示，可以通过四个步骤来描述。首先，内容必须在分布式系统中可用，也就是说，内容必须在标准服务(如发现和访问服务)中发布。其次，用户需要发现最终将通过使用这些服务访问的内容(第三步)。最后，用户处理内容并生成新的内容，这些内容应该在分布式系统关闭周期中集成和发布。

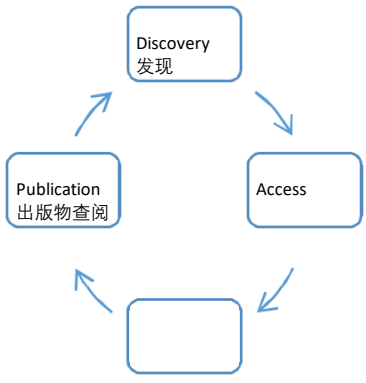

Figure 1. Content life cycle in Geographical Information Infrastructure[5]
图 1.地理信息基础设施的内容生命周期[5]

Nowadays, because of many regional and global initiatives on mapping, there are so many standard, and not every government or corporate comply with those standard. They usually developed the in-house applications with their own standard because of different data needs.
如今，由于许多地区和全球的测绘倡议，有这么多的标准，并不是每个政府或企业都遵守这些标准。由于不同的数据需求，他们通常使用自己的标准开发内部应用程序。

III. THE RESEARCH
研究

The concept of graphical and non-graphical data management had developed in the Laboratory of Computer Graphics and Artificial intelligence ITB, a NGGIS (Next Generation Graphical Information System) [6], which provide services of desktop-based GIS system that can use the data in a web network.
图形和非图形数据管理的概念是在计算机图形学和人工智能 ITB 实验室开发的，这是一个 NGGIS(下一代图形信息系统)[6]，它提供桌面地理信息系统服务，可以在网络中使用这些数据。

This research aims to develop spatial data management technology for the corporate model based on the Google Maps platform. The results of this study can be used as:
本研究旨在开发基于谷歌地图平台的企业模型空间数据管理技术。这项研究的结果可用于:

1. The basis for the management of corporate data visually-based on Google Maps
基于 GoogleMaps 的企业数据可视化管理的基础

   The model can be the basis of data management and general guidelines for various corporations, especially in Indonesia in visualizing its data based on Google Maps that can be useful for business development in Indonesia
   该模型可以作为数据管理的基础和各种公司的一般指导方针，尤其是在印度尼西亚，根据可用于印度尼西亚业务发展的谷歌地图可视化其数据

534

in the face of global competition, which in turn accelerate and expand economic development in I ndonesia.
面对全球化的竞争，印尼的经济发展进一步加速和扩大。

2. Liaison between corporate information systems with visualization of information.

企业信息系统与信息可视化之间的联络。

With the resulting technology, the information will be presented in a more intuitive and more attractive. Information can be presented visually and is positioned as a layer on Google Maps. With this approach, the analysis of visual information can be done so that the benefit of corporate interests
随着最终的技术，信息将以更直观和更有吸引力的方式呈现。信息可以直观地显示，并在 Google 地图上定位为一个层。利用这种方法，可以对视觉信息进行分析，从而使企业获得利益

Based on the purpose of the research and the result of the previous research, we modify the cycles proposed by Trilled et.al. (2013) to explain NGGIS data life cycle .
基于本文的研究目的和前人的研究成果，我们对 trillet.al 提出的循环进行了修正。(2013)解释 NGGIS 数据生命周期。