## Topic: Community Detection/Mining (CDM) of Road Network to enable business decisions

|  |  |  |
| --- | --- | --- |
| **Problem Definition** | **Potential Gaps In literature** | **Contribution** |
| Admin boundaries do not necessarily reflect natural communities | * + Barriers such as rivers, highway routes might divide a community   + Incorporation of public transit modelling uncommon   + Some studies incorporate POIs in classification of community but do not evaluate the importance/relevance of the POIs in the classification | Improvement of Community detection/Classification with additional features |
| Most research focuses on social network. Road network is much sparser than Social Network | * + Comparison of Algorithm in community detection for road network   + Comparison of CDM effectiveness in Road networks in Urban versus Rural, would they require different algo/parameters? | Framework to determine the right CDM for road network |
| Most research stops at clustering/classification of communities from CDM | * + Potential applications for business decision, e.g. Product assortment, location allocation of new businesses | Business applications |

## 20 to CiteSpace, v. 5.5.R2 (64-bit) November 1, 2019 PM CST WoS: onsipreparing\DOcurrentGlSbdata Timespan: 2010-2019 (Slice Length=3) Selection Criteria • Top 50 per slice, LRF=3.O, L BY—8 2 0 Netva.rk: N= 159. E=608 (Density=O.0484) Largest CC• 150 (94%) Nodes Labeled: 2.0% Pruning: None JANG 2010) GOODCHILD MF 7) #1 vgi HAKL•AYM 2 10) DCHILD MF (2012) AGE J (2009) #5 spatia econometrics 'CHANDER G (2009) GST KJ (20'3) hange detection U D (2007b (2011) p • NTIUS G (2011) FISHER PE f T 07 #7 laser scannin WAGNER W ( Review of Geospatial Publications in Past Decade

**Cluster #0 Landsat:**

e.g. “Object based image analysis for remote sensing” (Blaschke 2010 2010), Support vector machines in remote sensing: A review ” (Mountrakis et al. 20112011), and “ Death to Kappa: birth of quantity disagreement and

allocation disagreement for accuracy assessment ” (Pontius and Millones 2011 2011).

**Cluster #1, VGI (Volunteered Geographical Information),**

e.g. works include “ Citizens as sensors: the world of volunteered geography”(Goodchild 2007 2007), How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets datasets”(Haklay 2010 2010), and “Citizens as sensors: the world of volunteered geography ”(Goodchild and Li 2012 2012).

**Cluster #2 Social Sensing.**

“Social Sensing: A New Approach to Understanding Our Socioeconomic Environments” (Liu et al. 20152015), Discovering regions of different functions in a city using human mobility and POIs. POIs.”(Yuan et al. 20122012), and “Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends trends”(Weng 2012 2012).

**Cluster #3 Change detection**

E.g. “Assessing the accuracy of remotely sensed data data” in 2009, “ A survey of image classification methods and techniques for improving

classification performance performance”(Lu and Weng 2007 2007), and “Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors”(Chander et al. 20092009).

**Cluster #4 Spatial Econometric**

E.g. the representative works include “ A comparative assessment of decision-support tools for ecosystem services quantification and

valuation”(Bagstad et al. 20132013), “Introduction to spatial econometric”(LeSage and Pace 2009 2009),and Geospatial

Cyber infrastructure: Past, present and future” (Yang et al. 20102010).

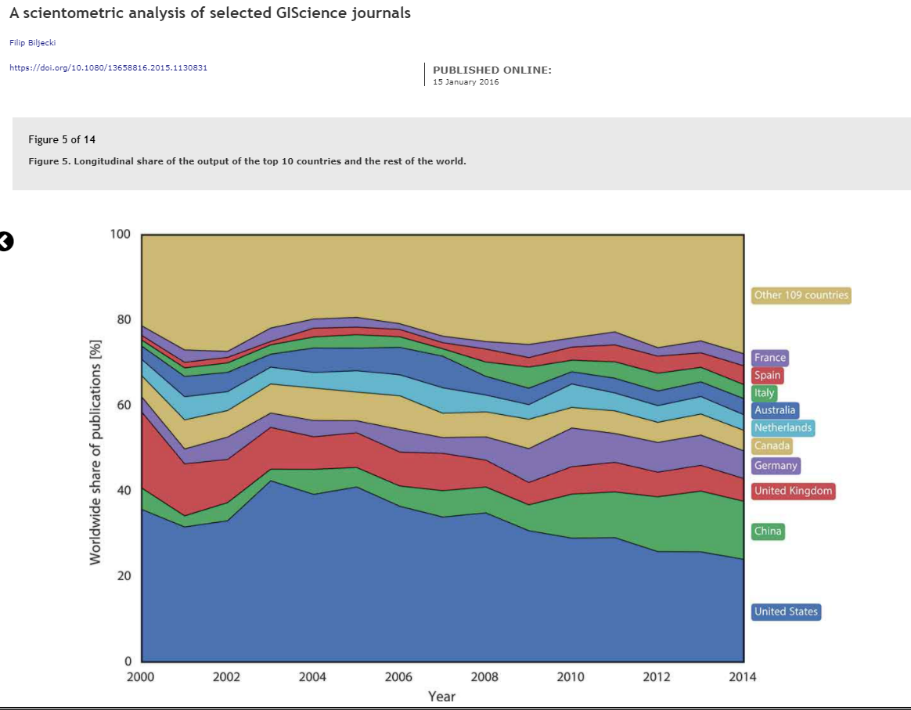
**Cluster #5 Laser Scanning**

E.g. “The shuttle radar topography mission” (Farr et al. 20072007), Gaussian decomposition and calibration of a novel small small-footprint full full-waveform digitising airborne laser scanner”(Wagner et al. 20062006), and “Causes and consequences of error in digital elevation models” (Fisher and Tate 2006 2006).

**New Topics not yet reflected in the clustering**

New topics like sensor networks, smart environments, 3D modeling, smart cities, IOT are also frequently discussed, but they are not well reflected in the clusters yet

Hu, K., Dai, Z., Zheng, J., & Wu, H. (2020). ANALYSIS OF THE STATUS AND TREND OF GIS RESEARCH IN THE PAST TEN YEARS FROM 2010 TO 2019. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B4-2020, 47–53. https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-47-2020



## Survey of Geospatial related Journals

|  |  |  |  |
| --- | --- | --- | --- |
| **Journal** | **Journal Description** | **Impact factor** | **Popular Articles** |
| **ISPRS journal of photogrammetry and remote sensing** | The Journal provides a channel of communication for scientists and professionals in all countries working in the many disciplines that employ photogrammetry, remote sensing, spatial information systems, computer vision, and related fields. The Journal is designed to serve as a source reference and archive of advancements in these disciplines.. | 7.32 | * Deep learning in remote sensing applications: A meta-analysis and review - Open access * Deep learning classifiers for hyperspectral imaging: A review - Open access * Object based image analysis for remote sensing - Open access |
| International journal of applied earth observation and geoinformation | The International Journal of Applied Earth Observation and Geoinformation publishes original papers that apply earth observation data to inventarisation and management of natural resources and the environment. | 4.65 | * The 2019 Brumadinho tailings dam collapse: Possible cause and impacts of the worst human and environmental disaster in Brazil * High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data - Open access * The urban greenness score: A satellite-based metric for multi-decadal characterization of urban land dynamics - Open access |
| Computers & geosciences | Geoscientific topics of interest include: mineralogy; petrology; geochemistry; geomorphology; paleontology; stratigraphy; structural geology; sedimentology; hydrology; hydrogeology; oceanography; atmospheric sciences; climatology; meteorology; geophysics; geomatics; seismology; geodesy; paleogeography; environmental science; soil science; glaciology. | 2.991 | * Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir area, China * Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data - * Displacement prediction of Baijiabao landslide based on empirical mode decomposition and long short-term memory neural network in Three Gorges area, China |
| International journal of geographical information science | Published research covers innovations in GIScience and novel applications of GIScience in natural resources, social systems and the built environment, as well as relevant developments in computer science, cartography, surveying, geography and engineering in both developed and developing countries. | 3.545 | * Delineating urban job-housing patterns at a parcel scale with street view imagery * Pattern-based identification and mapping of landscape types using multi-thematic data * Delineating urban job-housing patterns at a parcel scale with street view imagery |
| Annals of the association of american geographers | Articles are divided into four major areas: Geographic Methods; Human Geography; Nature and Society; and Physical Geography, Earth and Environmental Sciences. There are Editors responsible for each of these themes. | 3.302 | TBC |
| Computers environment and urban systems | Contributions emphasizing the development and enhancement of computer-based technologies for the analysis and modeling, policy formulation, planning, and management of environmental and urban systems that enhance sustainable futures are especially sought. The journal also encourages research on the modalities through which information and other computer-based technologies mold environmental and urban systems. | 4.655 | * A comparison between the cost effectiveness of CCTV and improved street lighting as a means of crime reduction - Open access * Social Media data: Challenges, opportunities and limitations in urban studies - Open access * A machine learning-based method for the large-scale evaluation of the qualities of the urban environment - Open access |
| Photogrammetric engineering and remote sensing | Photogrammetric Engineering & Remote Sensing commonly referred to as PE&RS, is the official journal of imaging and geospatial information science and technology. Included in the journal on a regular basis are highlight articles such as the popular columns “Grids & Datums” and “Mapping Matters” and peer reviewed technical papers. | 3.15 | * High frequency passive microwave radiometry over a snow-covered surface in Alaska * Microdensitometry to identify saline rangelands on 70-mm color-infrared aerial film |
| **ISPRS international journal of geo-information** | The journal covers original research in fundamental and computational geographic information science, including applying geographical information science to monitoring, prediction, and decision making, as well as natural resources, social systems, computer science, cartography, surveying, geography, and engineering, in both developed and developing countries. | 1.502 | * Spatial Distribution and Morphological Identification of Regional Urban Settlements Based on Road Intersections * Utilizing Urban Geospatial Data to Understand Heritage Attractiveness in Amsterdam * Multitemporal Analysis of Land Use and Land Cover within an Oil Block in the Ecuadorian Amazon |
| **Environment and Planning B: Urban Analytics and City Science** | leading journal for the publication of high-quality articles that present cutting-edge research in analytical methods for urban planning and design. The journal focuses on smart cities, urban analytics, GIS, and urban simulation models. It also deals with visualisation, computation, and formal design-based methods applicable to morphological processes and structures in cities and regions. | 2.822 | * Improving assessment of urban racial segregation by partitioning a region into racial enclaves * The geography of online shopping in China and its key drivers * Polycentric urban development in China: A multi-scale analysis * A multi-scale analysis of 27,000 urban street networks: Every US city, town, urbanized area, and Zillow neighborhood Geoff Boeing |
| International journal of digital earth | Develop methods that turn all forms of geo-referenced data, from scientific to social, into useful information that can be analyzed, visualized, and shared; Present innovative, operational applications and pilots of Digital Earth technologies at a local, national, regional, and global level; Expand the role of Digital Earth in the fields of Earth science, including climate change, adaptation and health related issues,natural disasters, new energy sources, agricultural and food security, and urban planning; Foster the use of web-based public-domain platforms, social networks, and location-based services for the sharing of digital data, models, and information about the virtual Earth; and Explore the role of social media and citizen-provided data in generating geo-referenced information in the spatial sciences and technologies. | 3.985 | * The characteristics of multi-source mobility datasets and how they reveal the luxury nature of social distancing in the U.S. during the COVID-19 pandemic * Spatial prediction of sparse events using a discrete global grid system; a case study of hate crimes in the USA * Antarctic-wide annual ice flow maps from Landsat 8 imagery between 2013 and 2019 |
| Transactions in GIS | GIS, GPS, Remote Sensing and related geospatial technologies; geospatial data acquisition and sensing; maps and spatial reasoning; spatial data infrastructures; standardization and interoperability; spatial data structures and databases; geocomputation; spatiotemporal analysis, integration and modeling; spatial data quality and uncertainty; GIS education and certification; GIS and society; location privacy; and desktop, mobile and Web-based spatially-enabled applications and services. | 2.119 | * An activity‐based framework for detecting human movement patterns in an urban environment * An analysis of the spatial and temporal distribution of large‐scale data production events in OpenStreetMap * Evaluating healthcare resource inequality in Beijing, China based on an improved spatial accessibility measurement |
| **GIScience & remote sensing** | GIScience & Remote Sensing publishes original, peer-reviewed articles associated with geographic information systems (GIS), remote sensing of the environment (including digital image processing), geocomputation, spatial data mining, and geographic environmental modelling. Papers reflecting both basic and applied research are published. | 3.588 | * A new type of dual-scale neighborhood based on vectorization for cellular automata models * Mapping rice area and yield in northeastern asia by incorporating a crop model with dense vegetation index profiles from a geostationary satellite * Recent land deformation detected by Sentinel-1A InSAR data (2016–2020) over Hanoi, Vietnam, and the relationship with groundwater level change |
| Cartography and geographic information science | Cartography and geographic information science | 2.271 | * Modeling the risk of robbery in the city of Tshwane, South Africa * Geo-located Twitter as proxy for global mobility patterns Bartosz Hawelka,Izabela Sitko,Euro Beinat,Stanislav Sobolevsky, |
| **Geographical analysis** | *Geographical Analysis* publishes geographical theory, model building, and quantitative methods to geographers and scholars in a wide spectrum of related fields. | 1.905 | * Delineating the Spatio‐Temporal Pattern of House Price Variation by Local Authority in England: 2009 to 2016 * Incorporating E‐commerce into Retail Location Model * Street Network Models and Indicators for Every Urban Area in the World Geoff Boeing |
| Geoinformatica | GeoInformatica presents the most innovative research results in the application of computer science applied to geographic information systems.  Publishes research at the intersection of computer science and geographic information science Covers spatial modeling and databases; human-computer interfaces for GIS; digital cartography; space imagery and more Chronicles progress toward more and more sophisticated computing tools in geographic studies | 2.81 | * Behavior-based location recommendation on location-based social networks * Learning evolving user’s behaviors on location-based social networks |
| Journal of geographical systems | Coverage includes regional science, economic geography, spatial economics, regional and urban economics, GIScience and GeoComputation, big data and machine learning. Spatial analysis, spatial econometrics and statistics are strongly represented.  One of the distinctive features of the journal is its concern for the interface between modeling, statistical techniques and spatial issues in a wide spectrum of related fields. An important goal of the journal is to encourage a spatial perspective in the social sciences that emphasizes geographical space as a relevant dimension to our understanding of socio-economic phenomena. | 1.324 | * Open source tools for geographic analysis in transport planning * Multidimensional spatiotemporal evolution detection on China’s rural poverty alleviation * Detecting space–time clusters of COVID-19 in Brazil: mortality, inequality, socioeconomic vulnerability, and the relative risk of the disease in Brazilian municipalities |
| Photogrammetrie fernerkundung geoinformation | Papers published in PFG highlight new developments and applications of these technologies in practice. The journal hence addresses both researchers and students of these disciplines at academic institutions and universities and the downstream users in both the private sector and public administration.  PFG places special editorial emphasis on the communication of new methodologies in data acquisition, new approaches to optimized processing and interpretation of all types of data which were acquired by photogrammetric methods, remote sensing, image processing and the computer-aided interpretation of such data in general. | 1.259 | * Supporting Management of Refugee Streams by Earth Observation and Geoinformation * Impact Assessment of Oil Exploitation in South Sudan using Multi-Temporal Landsat Imagery |
| Journal of spatial science | It covers cartography, geodesy, geographic information **science**, hydrography, digital image analysis and photogrammetry, remote sensing, surveying and related areas. | 1.711 | * Remote sensing of mangroves using unmanned aerial vehicles: current state and future directions * Urban slum detection using texture and spatial metrics derived from satellite imagery |
| Spatial cognition and computation | Spatial Cognition and Computation encourages the submission of articles on any topic in the areas of spatial cognition and spatial computation. Research articles, empirical studies, integrative reviews of the literature, and shorter opinion pieces will be considered for publication. Specific topics within the scope of the journal: cognitive and computational models of spatial cognition; navigation, environmental learning, and cognitive mapping; cognitive development and representation of large-scale space; common sense and qualitative models of space; robot navigation and wayfinding; models and applications of spatial and diagrammatic reasoning; visual and sign languages; visualization and spatial cognition; cognitive theories of imagery and imaginal reasoning; integration of vision and natural language processing; representation and processing of spatial expressions and crosslinguistic issues in spatial language; gestural analysis and multimodal interfaces; architecture and spatial cognition; philosophical issues in spatial cognition; and spatial cognition in virtual environments and hypermedia. | 2.808 | * Spatio-Temporal Analytics for Exploring Human Mobility Patterns and Urban Dynamics in the Mobile Age |

## CDM (Community Detection and Mining) Methodology

|  |  |
| --- | --- |
| **CDM Methodology** | **Extracted from Nawaz, W. (2019). *An Empirical Study of Community Detection Algorithms on Social and Road Networks*.** |
| **Clauset Newman Moore** | Clauset Newman Moore (CNM) [1] belongs to a hierarchical clustering strategy that maintains hierarchy of the resultant clusters. The hierarchical clustering method follows either agglomerative or divisive approach.CNM uses classical agglomeration approach that is a bottom up strategy while divisive method follows top down strategy. CNM optimizes the modularity of the nal partition by making greedy choices. This algorithm is effectively used for complex networks in research community, and was designed specically to analyze the community structures of extremely large networks, i.e. millions of nodes. |
| **Radicchi [18]** | Radicchi [18] is also a type of hierarchical clustering algorithm but unlike CNM, it takes a divisive approach by starting from the whole graph and splitting it into communities gradually. |
| **Label Propagation Algorithm (LPA)** | Label Propagation Algorithm (LPA) [7] is an ecient, near linear time, algorithm to detect community structures in large-scale networks. It is a semi-supervised algorithm that uses unlabeled nodes to finnd out the labels. It has an advantage in running time and performs well when there is prior information or annotated data. |
| **TopLeaders** | TopLeaders[4], i.e. Leadership expansion algorithm, extracts clusters from the graph identifying it as sets, consisting of a leader node and its follower nodes that are close to the leader. This algorithm requires to select initial k leaders as the number of desired communities. |
| **Sequential Clique Percolation** | Sequential Clique Percolation (SCP) [5] algorithm is based on the clique percolation method and detects k-clique subgrpahs for a given value of k from dense graph by sequentially inserting edges and keeping track of the emerging community structure. In comparison to CFinder[19], it finds all the cliques of single size and output the communities for all possible thresholds, while CFinder finds maximal cliques in a graph and produces communities of all possible clique sizes. Therefore, we can consider this algorithm as a alternative to CFinder. Another good thing about SCP algorithm is that it works well with large sparse graphs, however, it may not be a good option when the graph is very dense or it contains large size cliques. |
| **Matrix Blocking Dense Subgraph Extraction (MB-DSGE)** | Matrix Blocking Dense Subgraph Extraction (MB-DSGE) [2] algorithm reorders a relatively sparse graph and extracts dense subgraphs as communities. More precisely, for clustering, it constructs a hierarchy tree using matrix blocking technique, which groups similar columns of an adjacency matrix according to the cosine similarity measure. |
| **Walktrap** | Walktrap, developed by Pascal Pons, is an algorithm in graph theory, used to identify communities in large networks via random walks. These random walks are then used to compute distances between nodes. Nodes are then assigned into groups with small intra and larger inter-community distances via bottom-up hierarchical clustering. It should be noted, of course, that this algorithm considers only one community per node, which in some cases can be an incorrect hypothesis From <[*https://towardsdatascience.com/detecting-communities-in-a-language-co-occurrence-network-f6d9dfc70bab*](https://towardsdatascience.com/detecting-communities-in-a-language-co-occurrence-network-f6d9dfc70bab)> |
| **Infomap** | The basic idea behind the InfoMap algorithm is to use community partitions of the graph as a [Huffman code](https://en.wikipedia.org/wiki/Huffman_coding) that compresses the information about a random walker exploring your graph.  The central object is a random walker exploring the network with the probability that the walker transitions between two nodes given by its Markov transition matrix. At this point, we have effectively coded our network with an individual codeword for each node. However, in most real-world networks, we know that there are regions of the network such that once the random walker enters a region, it tends to stay there for a long time, and movements between the regions are relatively rare. This allows us to combinatorially combine codewords into Huffman codes: we can use a prefix code for each region, and then use a unique codeword for each node within a module, but we can reuse these node level codewords for each module. The same intuition can be gathered by looking at a street names; it would be crazy to have a unique street name for every street in the US, instead, we use states and towns, and then specify a street name, allowing us to reuse street names between towns (how many Main streets are there?).From <[*https://stackoverflow.com/questions/48528648/infomap-community-detection-understanding/54292999*](https://stackoverflow.com/questions/48528648/infomap-community-detection-understanding/54292999)> |