What will spatial analysis look like in twenty years?

By: Gaurav Bhardwaj

Advanced Spatial Analytics

Prof. Paul M. Torrens

Center for Urban Science & Progress, NYU

When we talk about future, we instantly think about the sci-fi realm, where we have floating cars, and everything is being done by the robots. Although it sounds very fascinating, we still have a long way to go to make our cities like the way we imagine. Also, with that another question arises, 'Do we want our cities to be such technologically advanced where human intervention is minimum? What will happen to humans if we have robots doing everything for us?'. That remains a very big challenge for decision makers to find a line where we need to stop relying on machines for everything. That being said, I definitely don't have prejudices against the technology and its use to build our modern-day cities. In fact, technology can facilitate a lot of innovations that would make human life much easier and less prone to errors. With that I'd like to start my discussion on 'what will spatial analysis look like in twenty years?'

With the advent of big data, we have seen innovations in almost every industry and those innovations have superseded what we thought was the extent of our capabilities. Spatial analysis, although a niche industry, has come a long way following the footsteps of big data innovations. We have researchers working on complex city planning and spatial analysis issues since early 1960s (Dempsey, 2014), but with the abundance of data to test the theories on, we can now be certain of what theories work for the present-day cities and what do not. One interesting articles I came across talks about the 10-year vision for the geospatial industry (Meier, 2016), and it had some of the ideas that I failed to account for when I first started brainstorming for this project. The notion of 'doing more for less' is more relevant than ever before. With data being publicly accessible and its integration with new technologies, one can achieve greater output with minimal supervision and manpower. The 'paradigm of geospatial information is changing', i.e. we are no longer limited to create maps and visualizations form geographic information, now this information could be applied with other data sources to add substantial meaning to the analysis. We are no longer constrained to tabular datasets for our analysis, and with integration of GIS data with computer vision capabilities, we can identify patterns and run analysis on the images taken from the satellite. Another advancement that we've seen in recent years is the 'incorporation of 3D and 4D technology' with the geographical datasets. This is particularly interesting because time domain adds the much-needed information that may be used for movement pattern analysis, emergency service response etc.

Although the future for GIS looks very promising from different aspects, I would like to concentrate on one such breakthrough that has revolutionized every industry it has touched. Artificial Intelligence, "is an entity (or collective set of cooperative entities), able to receive inputs from the environment, interpret and learn from such inputs, and exhibit related and flexible behaviors and actions that help the entity achieve a particular goal or objective over a period of time" (Daniel Faggella, 2017). This technology essentially trains a machine to do tasks that until now only human brain was capable of performing. With enough data and training, the models can surpass the human capabilities of prediction and classification. What has AI to do with GIS? With Machine Learning, a branch of Artificial Intelligence, we can train geo-spatial models to learn from the previous observations and re-train itself to accommodate the changes the model had initially missed. In one such article (Mansour Raad , 2017), the author discusses how ESRI made its way to AI by solving the traffic congestion problem for Kuwait City. They initially tried to code the defining rules of traffic but that was soon challenged by the exceptions such as the month of Ramadan or the surge of drivers on the first day of school. After multiple tweak-and-test, they realized that it would be difficult to address every scenario in code and still deliver an effective application. This was when they turned to ML for help, to automatically predict the traffic on roads given the defining parameters and past observations. They installed tracking devices on cars and analyzed the traffic patterns at particular hours during the day. This data was then fed to the model and the model predicted pretty accurate traffic congestion forecast, and also suggested new routes depending on the current location of the user.

That was just one application of leveraging geospatial AI to come up with better predictions. We covered one such model during our semester where we studied how geography can have an impact on general regression model. Regression models work a little different when dealing with spatial data, for one reason being that our assumptions that the observations should be independent of one another is violated (Charlton & Fotheringham, 2009). The authors suggest that if we plot the residuals on map, the neighboring spatial units will have similar magnitude and sign. These shortcomings of the traditional regression models are handled by Geographically Weighted Regression (GWR). In GWR, "an observation is weighted in accordance with its proximity to point *i* so that the weighting of an observation is no longer constant in the calibration but varies with *i*. Data from observations close to *i* are weighted more than data from observations farther away" (Fotheringham, Charlton, & Brunsdon, 1998). It essentially runs the OLS regression models for each of the spatial units and the observations are fitted to the function. The observations are given weights depending on their neighboring cells, i.e. for units closer to *i* will have higher weight than the units farther away.

Some other methods include using non-conventional datasets such as, using satellite imagery to make predictions. In 2010, when BP's deep-water drilling rig exploded in Gulf of Mexico, the company claimed that the spill was just 1,000 barrels per day. A small company in West Virginia studied the satellite data to analyze the truth behind the spill, and they concluded that the spill was 20 times bigger than claimed (Singh, 2017). In another attempt, UBS Investment Research team tried to predict the quarterly earnings of Wal-Mart by analyzing the number of cars in a parking lot of particular Wal-Mart store using satellite imagery. One of the most unique projects I recently came across was prediction of poverty using night-light satellite imagery for five African countries Nigeria, Tanzania, Uganda, Malawi, and Rwanda (Neal Jean, et al., 2016). The study by researchers in Stanford University predicted the poverty based on nighttime light as it is a rough proxy for economic wealth, and nighttime maps of the world show that many developing countries are sparsely illuminated. Since the imagery data is public and easily accessible, without any nonproprietary issue, the researchers were able to combine survey data to explain 75% of the variation in local-level economic outcomes. On similar notes, one of the startups in geospatial domain, Descartes Labs, has tons of high resolution satellite imagery which they use to predict food yields months in advance of current standards (Singh, 2017). This scalable and innovative approach has been widely appreciated and many government and for-profit firms have collaborated with them to alleviate food crisis even before it happens.

Another buzz word that has been catching on is Location Intelligence (LI). Location intelligence is the collection of insights we can gather from the interaction between people and physical locations (Burke, 2018). In future, advertisers and market analysts will depend on this technology to study the patterns in people's movement and their choice of selection. The GPS on our devices records the longitude and latitude of the location we walk through and this location information can be used by the market analysts to see which part of the city we spend our time in, visit what kind of store and indulge in what kind of activities. As creepy as it sounds, that is the power of location intelligence in retail marketing and advertising. Without knowing someone in person, we now have the ability to know almost everything about their behavior. Let's discuss some of the advantages of using this technology as described by Melissa Burke in one of her articles (Burke, 2018). By analyzing the location of users in real time, Skyhook analyzed the movement of spectators at the U.S. Open and was able to display the most popular viewing spots, escape strategy, after the game was over and in general how people interact with the space (Sharma, 2017). The fractal dimensions and space-time geography that we studied in class can come handy here since we'll be dealing with complex path movement and these movements can be visualized with the time dimension to further enhance our analysis. During the 2017 Super Bowl, Skyhook tried to capture the preferences of Falcon and Patriots fans so see if there are any distinctions in their movements and habits. Analyzing where people go on a regular day, or even on any event day, help you get insights to engage with them effectively through marketing campaigns in a more personalized way (Morra, 2017). One of the applications that I found interesting from retail marketing perspective is geo-conquesting. LI allows you to create a geo-fence over the stores/space of your rivals to get a better understanding of foot traffic, buying preferences etc. (Burke, 2018). These insights could help us come up with marketing strategies to lure the customers from rival stores by offering better discounts and privileges.

While we talk about all the innovative and interesting things we can do with spatial data, we still haven't discussed how all this data will be processed. We had all this data for years but the ability to analyze such data on our computers had been a challenge. With ever increase in computing capabilities, especially after Big Data, GIS industry has also jumped into the research of finding fastest ways to process the tasks. We can imagine working on geographical dataset having millions of data points, and also the time and computation required to perform even the smallest operation. Geospatial Computing term was first explicitly coined Max J. Egenhofer paper, where he talks about Geo Sketch Pads, Smart Compasses, and Smart Glasses as part of "a vision of ubiquitous geospatial computing" (Egenhofer, 1999). With advancements in technology, we've resorted to high performance machines to get results in reasonable amount of time. Some of the advancements we've witnessed in geospatial industry is the usage of General Purpose Graphic Processing Unit (GPGPU), which provides an alternative and complementary solution to an existing cluster based high-performance geospatial computing. In one of the research papers (Zhang & You, 2013), the authors propose a practical approach to simplifying high-performance geospatial computing on GPGPUs by using parallel primitives. The authors define parallel primitives as a collection of fundamental algorithms that can be run on parallel machines. While generic parallel primitives have been applied to quite a few domains, in this paper the authors explain its application on geospatial indexing. They re-implemented BMMQ-Tree construction algorithm with their method and witnessed 10x improvement in time taken while reducing the coding complexity.

When talking about future of spatial analysis, I cannot not talk about a paper I read a few months back (?). 'The future of Geospatial Intelligence' (Dold & Groopman, 2017) talks about 'perceptality'- where geospatial perception meets reality, and the future of geospatial intelligence. The paper essentially talks about the applications of sensors, robotics, cameras, ML amongst others and how they contribute towards perceiving and capturing reality. The authors define geospatial intelligence as "the discipline inherent in capturing the physical dimension of the intersection between Information Technology (IT) & Operational Technology (OT) is the field of geospatial intelligence". This would include perception, cognition, computation, control, reaction, and understanding of physical features and geographically referenced activities (Dold & Groopman, 2017). All the data that we generate today is stored on the cloud, which had given a new face to 'client-server model' by creating a network between the distributed systems. Although it is great, there are some processes where we want low latency, for instance in healthcare services where the vitals of patients may change drastically. Any delay in information retrieval of this information can be catastrophic for the patient. The cloud server would take some amount of time to extract that information from the silos of data stored. Imagine a case of autonomous vehicles, where the information that is being recorded would be needed back in split seconds to perform further operations, e.g. taking a turn from a current path requires location information of the car as well as the surrounding. If the response from cloud takes more time than anticipated, we could get into an accident. For such scenarios, where low latency is priority, there's another technology called 'edge-computing'. What edge computing essentially does is it stores the information locally for a period of time before pushing it to the cloud network (Brandon Butler, 2017). It works as semi-permeable layer between the IoT device and the cloud server where data is held temporarily and the speed to information retrieval is high because of low data volume.

In this paper, I've discussed a few advancements that we're currently witnessing in geospatial industry and these technologies will only grow with time and further innovations. With implementation of deep learning with motion sensors or imagery data will further enhance our abilities to extract relevant information from the data and provide better visualizations of what happens in reality. The future cities may or may not be full of hovercrafts but the innovations in remote sensing techniques would definitely change the face of geospatial studies. With sophisticated datasets we would also require equally sophisticated hardware tools such as GPU and cloud platforms to network the distributed computing and be able to share the data and insights we generate. One of the elements I didn't mention explicitly above, but is equally ever-growing, is Agent Based Modelling (ABM). It is a software model where agents interact with each other and these interactions produces outcomes based on the rules we've established. The distinguishing feature about ABM is that the interactions are governed by the rules of the behavior that modeler simply encodes directly into the individuals who populate the environment (Chris House, 2014). With that addition, I'd like to conclude by saying that the more microscopic data we generate, the more computational capabilities will have to be developed, and the more accurate our models will be.

Bibliography

Brandon Butler. (2017, 09 21). *What is edge computing and how it's changing the network*. Retrieved from Network World: https://www.networkworld.com/article/3224893/internet-of-things/what-is-edge-computing-and-how-it-s-changing-the-network.html

Burke, M. (2018, 03 08). *The Nuts and Bolts of Location Intelligence*. Retrieved from Skyhook Wireless: https://www.skyhookwireless.com/blog/nuts-and-bolts-of-location-intelligence

Charlton, M., & Fotheringham, A. (2009, 03 03). *GEOGRAPHICALLY WEIGHTED REGRESSION .* Retrieved from http://gwr.maynoothuniversity.ie/wp-content/uploads/2016/01/GWR_WhitePaper.pdf

Chris House. (2014, 08 12). *Are Agent-Based Models the Future of Macroeconomics?* Retrieved from Orderstatistic: https://orderstatistic.wordpress.com/2014/08/12/are-agent-based-models-the-future-of-macroeconomics/

Dangermond, J. (2014, 12). *Geospatial Technology and the Future of the City*. Retrieved from ESRI ArcNews: http://www.esri.com/esri-news/arcnews/winter1415articles/geospatial-technology-and-the-future-of-the-city

Daniel Faggella. (2017, 05 15). *What is Artificial Intelligence? An Informed Definition*. Retrieved from TechEmergence: https://www.techemergence.com/what-is-artificial-intelligence-an-informed-definition/

Dempsey, C. (2014, 02 11). *GIS Timeline*. Retrieved from GIS Lounge: https://www.gislounge.com/gis-timeline/

Devia, N., & Weber, R. (2013). Generating crime data using agent-based simulation. *Computers, Environment and Urban Systems*, 26-41.

Dold, J., & Groopman, J. (2017, 05 16). *The future of geospatial intelligence* . Retrieved from Taylor & Francis Online: https://www.tandfonline.com/doi/full/10.1080/10095020.2017.1337318

Egenhofer, M. J. (1999). *Spatial Information Appliances: A Next Generation of Geographic Information Systems.* Retrieved from Divisão de Processamento de Imagens: http://www.dpi.inpe.br/geoinfo/portuguese/geoinfo1999/papers/Max.pdf

Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A, 30*, 1907-1908.

Goodchild, M. F. (2010). Twenty years of progress: GIScience in 2010. *JOURNAL OF SPATIAL INFORMATION SCIENCE*, 3-20.

Mansour Raad . (2017, 04 18). *A new business intelligence emerges: Geo.AI*. Retrieved from ESRI: https://www.esri.com/about/newsroom/publications/wherenext/new-business-intelligence-emerges-geo-ai/

Meier, P. (2016, 03 14). *A 10 Year Vision: Future Trends in Geospatial Information Management*. Retrieved from iRevolutions: https://irevolutions.org/2016/03/14/future-trends-in-geospatial/

Morra, R. B. (2017, 02 03). *Super Bowl LI: Brand Preferences & Behaviors of Patriots & Falcons Fans*. Retrieved from Skyhook Wireless: https://www.skyhookwireless.com/blog/super-bowl-2017-patriots-falcons-fan-behavior

Neal Jean, Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016, 08 19). *Combining satellite imagery and machine learning to predict poverty*. Retrieved from American Association for the Advancement of Science: http://science.sciencemag.org/content/353/6301/790

Sharma, K. (2017, 06 19). *Skyhook's Location Intelligence at the 2017 USGA Golf Innovation Experience*. Retrieved from Skyhook Wireless: https://www.skyhookwireless.com/blog/apps/skyhook_2017_usga_golf_innovation_experience

Singh, I. (2017, 11 14). *The perfect storm called artificial intelligence and geospatial big data*. Retrieved from Geo Awesomeness: http://geoawesomeness.com/the-perfect-storm-called-artificial-intelligence-and-geospatial-big-data/

Tzafestas, S. G., Laliotis, L. N., & Protonotarios, M. (2017, 05 26). *Decision Support and Artificial Intelligence in GIS: Overview and Applications.* Retrieved from ScienceDirect: https://www.sciencedirect.com/science/article/pii/S147466701736041X

Zhang, J., & You, S. (2013, 06 01). *Simplifying High-Performance Geospatial Computing on GPGPUs Using Parallel Primitives: A Case Study of Quadtree Constructions on Large-Scale Geospatial Rasters .* Retrieved from Semantic Scholar: https://pdfs.semanticscholar.org/8765/634e88178ffd8d7c4de770b89c00d644340c.pdf