Location Powers: Data Science

Session 1 Summary Foundations

George Percivall
OGC

Nov 13&14, 2019

OGC®

Location Powers: Data Science

Session 1: Foundations

- Moderator: Ed Parsons, Google
- Rapporteur: George Percivall, OGC

- Marc Armstrong - Presentation:
  - Fundamental Issues in Geospatial Data Science: Emerging Trends in Data and Analytics
- Nils Lahr - Presentation:
  - Scaling machine learning to handle visual data will result in more powerful AI

- Panel on Foundations/Motivations :
  - Kathleen Stewart, UMCP/CGIS - New opportunities through big mobility data analytics
  - Anand Padmanabhan, University of Illinois
  - Mark Korver, AWS
  - Jayant Sharma, Oracle

Session 1. Foundations – **1. What Works**

- Massive quantities of geospatial data are now being streamed in real time by, for example, IoT sensors, telematics, UAVs and CubeSats.
- Parallel analysis of geospatial data - particularly those employing distributed cyberinfrastructure - parallelism can address some volume issues
- Video Analytics, Machine Learning, and Big Data Analytics are all coming together to create a new form of Artificial Intelligence.
- Increasing availability of smartphones and location-based services apps, massive mobile device location data is being generated containing detailed information about trips and travel routes.

Session 1. Foundations – **2. Open Questions**

- In the past, most applications were developed assuming static information or “data at rest”. This is increasingly not the best option.
- data steam velocities overwhelm those methods that have been developed over decades to populate conventional GIS toolboxes.
- To address the volume, veracity, and live-streams of visual data, the industry is having to abandon its mature structured data solutions.
- Being able to extract activities and actions from multiple sensor types, such as cameras and IoT devices, will allow us to build accurate digital-twins or a virtual representation of the real world.

Session 1. Foundations – **3. Next Steps**

- Entirely new classes of sampling strategies and analytical methods must be developed to support the geospatial analysis of high-velocity data streams.
- To build these models, we need more innovation around how to map the output of visual machine learning models back into the temporal and spatial domains.
- With accurate ‘twins’, we can start to predict the future, understand the world at a macro level, and empower humans with superhuman abilities
- Deeper understanding of mobility: challenges in harnessing mobile and LBS data and extracting meaningful information about mobility patterns

FUNDAMENTAL ISSUES IN GEOSPATIAL DATA SCIENCE: EMERGING TRENDS IN DATA AND ANALYTICS

Marc P. Armstrong, The University of Iowa, marc-armstrong@uiowa.edu

- Advances in geospatial data capture technologies are fundamentally changing the way that GIS applications are conceptualized. In the past, most applications were developed assuming static information or “data at rest”. This is increasingly not the best option.
- Massive quantities of geospatial data are now being streamed in real time by, for example, IoT sensors, telematics, UAVs and CubeSats. Taken alone, the volume of this data poses computational challenges for many types of analyses.
- While that is bad enough, data steam velocities overwhelm those methods that have been developed over decades to populate conventional GIS toolboxes. Even worse, these data problems are occurring when hardware performance has begun to plateau with the suspension of Moore’s Law.
- In this presentation I discuss emerging streaming data sources and briefly review approaches to parallel analysis of geospatial data, particularly those employing distributed cyberinfrastructure. Though parallelism can address some volume issues, entirely new classes of sampling strategies and analytical methods must be developed to support the geospatial analysis of high-velocity data streams.
- These new types of analyses must be developed with an eye on emerging hardware trends including neuromorphic and heterogeneous architectures. And then, of course, there is the polarizing quantum quandary. But there is good news: there are plenty of interesting challenges ahead.

SCALING MACHINE LEARNING TO HANDLE VISUAL DATA WILL RESULT IN MORE POWERFUL AI

Nils Lahr, CEO Orions Systems

- Video Analytics, Machine Learning, and Big Data Analytics are all coming together to create a new form of Artificial Intelligence. In order to address the volume, veracity, and live-streams of visual data, the industry is having to abandon its mature structured data solutions.
- Additionally, unlike abstract financial and text-based data, “visual” data contains information about real-world activities. Being able to extract activities and actions from multiple sensor types, such as cameras and IoT devices, will allow us to build accurate digital-twins or a virtual representation of the real world.
- To build these models, we need more innovation around how to map the output of visual machine learning models back into the temporal and spatial domains.
- Once we have such accurate ‘twins’, we can start to predict the future, understand the world at a macro level, and empower humans with superhuman abilities.

NEW OPPORTUNITIES THROUGH BIG MOBILITY DATA ANALYTICS

Kathleen Stewart, Center for Geospatial Information Science, Univ. of Maryland

- We all travel a great deal, whether it is in the context of our families and social ties, or for our jobs. Our daily choices regarding where we go and how we get there, contribute to a much larger picture of daily travel patterns.
- Much of our movement is routine and regular but there are also occasions when travel patterns vary, for example, for a major sporting event or during a flood or hurricane hazards when evacuations occur.
- With the increasing availability of smartphones and location-based services apps, massive mobile device location data is being generated containing detailed information about trips and travel routes.
- These data offer great potential for providing insights and a deeper understanding of mobility. My panel talk will consider these emerging and rapidly expanding data sources and what challenges lie in harnessing these data and extracting meaningful information about mobility patterns.