geoEdge: A Real-time Analytics Framework for Geospatial Applications

Ranga Raju Vatsavai*
North Carolina State University
Raleigh, North Carolina
rrvatsav@ncsu.edu

Bharathkumar Ramachandra
North Carolina State University
Raleigh, North Carolina
bramach2@ncsu.edu

Zexi Chen
North Carolina State University
Raleigh, North Carolina
zchen22@ncsu.edu

John Jernigan
North Carolina State University
Raleigh, North Carolina
jajerni2@ncsu.edu

ABSTRACT

In many real-world applications, data looses its value if its not analyzed in real-time. Examples include natural disasters, crop disease identification and bioterrorism, traffic monitoring, monitoring human activities and public places, gas pipeline monitoring for leaks. Edge computing refers to pushing computing power to the edge of the network or bringing it closer to the sensors. We envision that an integrated framework (sensors + edge computers + analytics) allows near realtime analytics at the edge, which is critical for first responders to national security agencies alike. In addition to the generation of real-time actionable knowledge, edge computing allows compressing/reducing big geospatial data that need to be transmitted to centralized cloud or data centers. In this study, we present the vision behind geoEdge, and show feasibility results using feature extraction and unsupervised learning on an edge computing device.

CCS CONCEPTS

• **Computing methodologies → Distributed computing methodologies**; *Machine learning algorithms*; • **Computer systems organization → Embedded systems**.

KEYWORDS

real-time actionable knowledge, edge computing, remote sensing

ACM Reference Format:
Ranga Raju Vatsavai, Bharathkumar Ramachandra, Zexi Chen, and John Jernigan. 2019. geoEdge: A Real-time Analytics Framework for Geospatial Applications. In *Proceedings of BigSpatial '19*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3356999.3365468

*Corresponding Author; Affiliated with the Center for Geospatial Analytics.

BigSpatial '19, November 5, 2019, Chicago, IL, USA

ACM ISBN 978-1-4503-6966-4/19/11…$15.00
https://doi.org/10.1145/3356999.3365468

1 INTRODUCTION

A recent Computing Community Consortium (CCC) visioning activity defined spatial computing as a unifying field that "encompasses the ideas, solutions, tools, technologies, and systems that transform our lives by creating a new understanding of locations – how we know, communicate, and visualize our relationship to locations and how we navigate through them." [9]. Advances in sensor technologies have greatly facilitated collection and archival of big spatial and temporal data, that lead to the centralized processing. Planet Labs is collecting about 9 PB/year and DigitalGlobe (Maxar) is collecting 36 PB/year of remote sensing data alone. In a recent study, IDC [4] predicted that the global data will grow from 33 Zettabytes in 2018 to 175 Zettabytes in 2025. Using commodity 8 TB disk drives, we need more than 125 million hard disks to hold a Zettabyte of data, making offline analytics infeasible without employing smart data and edge computing technologies. In addition, there are many real-world applications, such as, natural disasters, crop disease identification and bioterrorism, traffic monitoring, monitoring human activities and public places, gas pipeline monitoring for leaks, and autonomous vehicles, where real-time extraction of knowledge from these data streams becomes critical. We envision that the edge computing frameworks that bring computing closer to the sensors (e.g., UAVs) is very relevant for the spatial computing community. Figure 1 shows an example architecture of spatiotemporal edge computing. This figure also shows the distinction between traditional off-line (backend) analytics and online (near) real-time analytics.

1.1 Why Now?

We see two key technology innovations that is driving this new vision. First, we observe that edge computing is not a new technology. Though ideas were in development for sometime, due to limited computing capabilities and power requirements, deploying computing closer to the field sensors was prohibitive. Second, unlike traditional sensors, the data generated by modern remote sensing sensors (e.g., UAV based) and in-situ sensors (e.g., optical soil moisture sensors) poses problems in terms of data transmission and storage. However, with recent advances in computing, especially embedded supercomputing chips (e.g., Nvidia's Jetson TX-1, TX-2, Nano), tiny yet powerful edge computers (e.g., Lenovo's P330 Tiny), and high-end edge workstations (e.g., Nvidia's EGX), spatiotemporal edge computing at the edge can be feasible and impactful.

As per Gartner, "91% of today's data is created and processed in centralized data centers. However, by 2022 about 75% of all data will need analysis and action at the edge." We now briefly describe the enabling technologies behind this vision and showcase few important applications, and argue why spatial computing community should pay attention to this emerging theme.

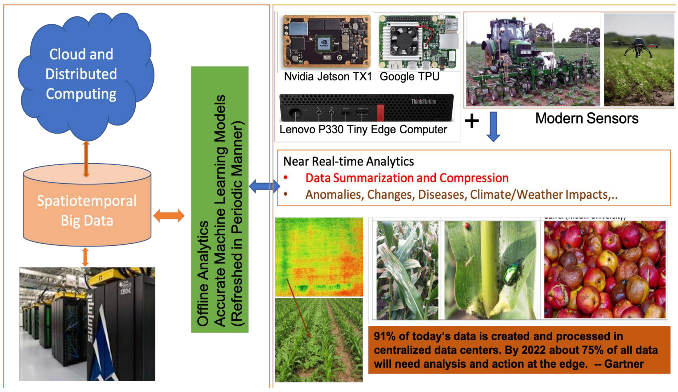

Figure 1: Traditional Offline Analytics Vs. Edge Based Near Real-time Analytics

2 THE RISE OF THE SENSORS

Gartner estimates that there will be 25 billion things connected to the Internet by 2020[1]. These sensors (or things) range from simple temperature sensors to complex ultrasound sensors. They are placed at fixed locations, or on moving platforms, or on remote platforms like traditional satellites to recent UAVs. Most of these sensors are connected to the Internet to transmit data to centralized facilities, including traditional data centers or modern cloud storage. In particular, we are interested in multi-spectral and hyper-spectral sensors mounted on UAVs which allows us to monitor natural resources and critical infrastructures (e.g., crops, waters, forest, cities, nuclear facilities). One major limitation with current centralized storage and offline data processing is that this infrastructure can neither support nor scale well for near realtime applications. A large portion of the data generated from these sensors is spatiotemporal in nature. Spatiotemporal data is often big, therefore it make sense to process this data close to where it is being generated, which makes case for *edge computing*.

3 THE RISE OF EDGE COMPUTING

The need for edge computing is evident from recent report "IDC FutureScape: Worldwide Internet of Things 2017 Predictions." As per this report, "By 2019, at Least 40% of IoT-Created Data Will Be Stored, Processed, Analyzed, and Acted Upon Close to, or at the Edge of, the Network." There are two key requirements for edge computing to be successful, (i) computing devices with low energy requirements, and (ii) highly scalable streaming analytics platforms.

[1]http://www.gartner.com/technology/research/internet-of-things/

4 THE RISE OF THE LOW ENERGY HIGH PERFORMANCE COMPUTING (AKA EMBEDDED SUPERCOMPUTING)

The power and efficiency of GPUs has increased drastically in the past decade [8] along with ever-cheaper and faster storage. Only 20 years ago, the fastest computer in the world achieved a performance metric of 1 teraFLOP [10] with energy consumption of 850,000 watts. The current NVIDIA Jetson TX1 platform can achieve the same performance metric at under 15 watts, and weighs 88 grams. The progress is absolutely astonishing. We effectively can put supercomputers in UAVs, but the algorithms and software have not kept pace. Fortunately, we are increasingly seeing algorithms adapted for streaming GPU-processed workflows.

5 THE RISE REAL TIME SPATIOTEMPORAL APPLICATIONS

We now describe two real world applications making the case for spatiotemporal edge computing. More comprehensive with geospatial big data, applications, and analytics can be found in [1, 3, 11, 12]

Towards UAV-based GPU-accelerated Remote Sensing: In aerial remote sensing, such as with unmanned aerial vehicles, edge computing with embedded GPUs can leverage immense computational power and efficiency [8] to provide actionable insight in the food-energy-water (FEW) nexus. NSF and several other federal agencies are considering FEW [2] to be next big research thrust in the USA. Food, energy, and water are essential for human wellbeing. Agriculture accounts for 70% of total freshwater withdrawals. About the 30% of global energy is consumed in food production. Figure 2 shows simple interaction between these three systems, where energy being is used to pump water for crops. In this context, spatiotemporal edge computing can be used to analyze data from UAVs to estimate in near real-time the soil moisture and crop water stress at field scale which allows farmers to schedule irrigation more optimally.

Figure 2: Image shows nexus between food, energy, and water systems

[2]https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505241

Farmers and agricultural producers cannot wait very long for remote sensing data to process in the cloud when an immediate decision must be reached regarding water application to a crop field, or how much fertilizer to use (and where to apply it) to address nutrient deficiency. In search and rescue operations, the immediacy of data processing is obvious. Embedded supercomputers (or GPUs) can greatly assist with remote sensing algorithms due to the inherent multidimensionality of the image data and the processors' highly parallel computational architecture [5] [8]. The effect would be to expedite the remaining processing, if not obviate some offline processing outright. The following are some example cases:

(1) Batch-to-stream workflows: In some cases, what has been an offline batch process like orthorectification of remote sensing images (image stitching) can happen in real time if hardware can leverage efficient parallel processing [2] [13].
(2) Anomaly detection: traditional streaming adaptations such as using a sliding window on real time data can allow for streaming anomaly detection with GPU acceleration and modified algorithms [13]. Rapidly processed anomaly detection could be of great benefit when assessing crop damage due to extreme weather events such as floods or large hail.
(3) Image classification: Multispectral or hyperspectral imagery can allow detection of crop pestilence or disease using methods such as support vector machines [7].
(4) Hyperspectral signature classification: Water contamination could be detected from hyperspectral data by flying UAVs along waterways. On-board processing for certain signatures and thresholds could trigger a flight reassignment to gather higher resolution imagery in suspected areas for offline processing. This approach could lead to more regular monitoring of waterways for threats to healthcare, agriculture, and the ecosystem.
(5) Data compression and summarization: Data compression and summarization (e.g., clustering) are efficient techniques to reduce data storage and transmission costs. As these are unsupervised techniques, it is easy to implement these algorithm on edge computers and compress images on the fly.

Figure 3 shows an use case of near real-time detection of weeds using the spatiotemporal edge computing framework. Such real-time detection of weeds and crop diseases will allow timely action by the farmer before its too late.

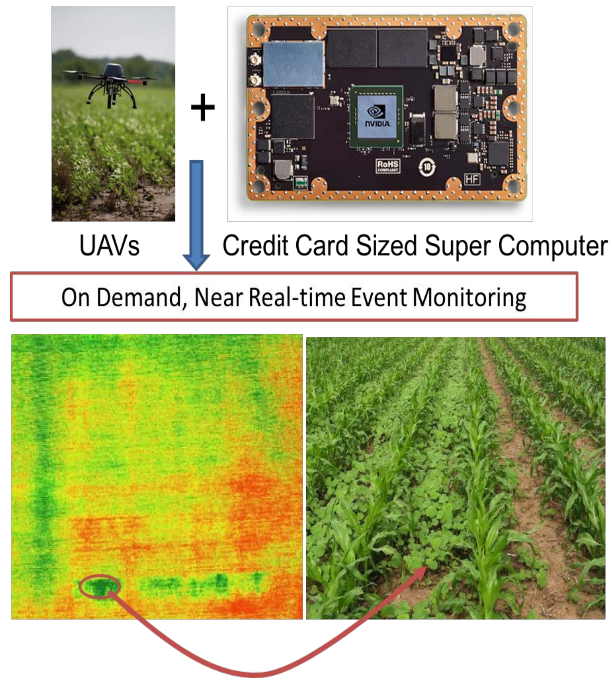

Figure 3: Near real-time weed detection while the data still being collected by UAVs

Anomaly Detection in Video Streams: Anomaly detection in video streams have several security applications. For example, one could monitor video streams at airports and other public places to identify unattended bags or other objects. We recently developed a probabilistic anomaly detection framework (KDD-17 under review). We ran this algorithm on a video dataset "Peds1" [6] consisting of natural images of a pedestrian walkway. Given that we model variation in local spatio-temporal neighborhoods, we would expect the method to behave similar to a 3D feature (optical flow) detector. Anomalous events could be abnormal motion patterns or high amount of variation in pixel values that lasts for a short amount of time. Figure 4 shows the result of applying our anomaly detection method to a video dataset. (a) represents the frame on which we queried for anomalies with respect to the rest of the video. Darker regions in (b) represent low cumulative chi-squared scores and thus low deviation and whiter regions represent high deviation from normality (e.g., person on bicycle on a pedestrian pathway). Finding such anomalous patterns in realtime is highly useful.

Figure 4: Anomalous events detected for a frame of a video of a pedestrian walkway

6 FEASIBILITY STUDIES

To study the feasibility of running complex models on the edge, we started with benchmarking the running times for widely used unsupervised approaches such as feature extraction (e.g., vegetation indices) and data summarization (e.g., clustering) methods. We used Lenovo ThinkStation P320 Tiny edge computer equipped with an Nvidia Quadro P600 GPU for these experiments.

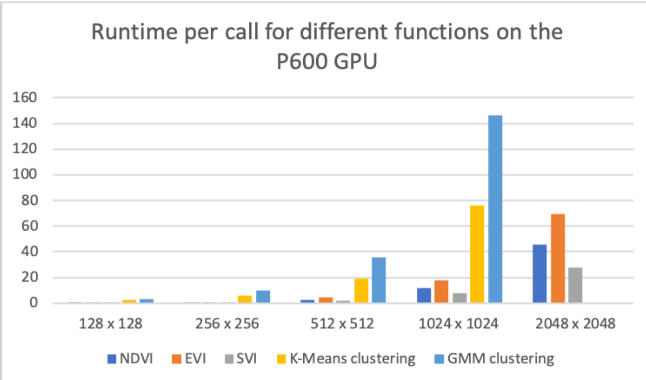

Figure 5

Experiment details: For these experiments we considered (i) Normalized Difference Vegetation Index, (NDVI) (ii) Enhanced Vegetation Index (EVI), (iii) Standardized Vegetation Index (SVI), (iv) K-means clustering and (v) Gaussian Mixture Model (GMM) clustering. For the clustering experiments, a 10% sample of input pixels is used to build the clustering model (with 10 clusters) and every pixel is classified with this model. K-means clustering used the smaller of 1000 iterations or a difference of 1e-2 on root mean squared error as the stopping criterion. GMM clustering used the smaller of 1000 iterations or a difference of 1e-8 on total log likelihood between successive iterations as the stopping criterion. To implement these functions we used python3 with the pytorch GPU acceleration library.

For image input, we synthetically generate an image by picking values in [0, 255] from a uniform distribution. We ran each of of these functions 1000 times and compute the average runtime per function call to summarize scalability.

Results: Figure 5 shows how the runtime scaling as function of image size. Since SVI, NDVI and EVI use 1, 2 and 3 bands to perform their computation and the data transfer to GPU memory is the most expensive piece, we see this trend reflected in their runtimes. We see that with even slightly more complex tasks such as K-means clustering or GMM clustering, the runtime per call quickly approaches 100 ms, and depending on the frequency of input images, may not meet a real-time processing requirement on the edge for large image sizes (I/O bound) and compute intensive tasks. Moreover, when the image size is too large (2048 x 2048), even a single image is too large to fit into GPU memory, throwing an OutOfMemory error. We are also working on parallelizing these algorithms on the Jetson TX-1/2 cards. We are also working with our UAS collaborators to integrate Jetson TX-1 for onboard processing of image streams.

7 CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

Now we have all ingredients required to enable *spatiotemporal computing at the edge*. If we suppose that a large problem in the food-energy-water nexus is inefficient agricultural resource use and a solution is advanced decision support, then a missing link in the chain between the two is a suite of algorithms and software to take advantage of edge-based supercomputing with efficient GPU (and hybrid GPU-CPU) systems such as the NVIDIA Jetson series, or even FPGA-CPU hybrids embed on UAVs. To conclude, we should state that for effective computation at the edge and achieving real-time performance with more complex tasks, the data transfer between CPU RAM and GPU DRAM should happen as the computation is being performed, asynchronously. Other factors such as size of GPU memory and optimizing computations should also be given a careful thought.

In order for edge-based spatial computing to have an impact on real-world problems, experts from multiple disciplines must collaborate to identify and develop right solutions. We believe that the time is ripe to form a community around the topic of geospatial edge computing and attack major challenges of spatiotemporal edge computing.

8 ACKNOWLEDGEMENTS

We would like to thank our collaborators, Rada Chirkova, and STAC lab members at NCSU, and Jaime E. Puente, Cassidy Lammers and Mike Leach at Lenovo. We greatly appreciate the research funding provided by the NSF funded IUCRC Center for Accelerated Real Time Analytics (CARTA) and Lenovo for donating ThinkStation P320 Tiny edge computer.

REFERENCES

[1] Budhendra L. Bhaduri, Dilip R. Patlolla, Ranga Raju Vatsavai, Anil M. Cheriyadat, Wei Lu, and Rajasekar Karthik. 2014. Emerging trends in monitoring landscapes and energy infrastructures with big spatial data. *SIGSPATIAL Special* 6, 3 (2014), 35–45.
[2] Dai Chenguang and Yang Jingyu. 2011. Research on orthorectification of remote sensing images using GPU-CPU cooperative processing. In *Image and Data Fusion (ISIDF), 2011 International Symposium on*. IEEE, 1–4.
[3] Arie Croitoru, rew Crooks, Jacek Radzikowski, Anthony Stefanidis, Ranga Raju Vatsavai, and N. Wayant. 2014. Geoinformatics and Social Media: New Big Data Challenge.
[4] John Rydning David Reinsel, John Gantz. The Digitization of the World, From Edge to Core. https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf. (????). [Online].
[5] Yan Ma, Lajiao Chen, Peng Liu, and Ke Lu. 2016. Parallel programing templates for remote sensing image processing on GPU architectures: design and implementation. *Computing* 98, 1-2 (2016), 7–33.
[6] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos. 2010. Anomaly detection in crowded scenes. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1975–1981. https://doi.org/10.1109/CVPR.2010.5539872
[7] T Rumpf, A-K Mahlein, U Steiner, E-C Oerke, H-W Dehne, and L Plümer. 2010. Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture* 74, 1 (2010), 91–99.
[8] Javier Setoain, Manuel Prieto, Christian Tenllado, and Francisco Tirado. 2008. GPU for parallel on-board hyperspectral image processing. *The International Journal of High Performance Computing Applications* 22, 4 (2008), 424–437.
[9] Shashi Shekhar, Steven K. Feiner, and Walid G. Aref. 2015. Spatial Computing. *Commun. ACM* 59, 1 (Dec. 2015), 72–81. https://doi.org/10.1145/2756547
[10] Joel O Stevenson, Robert A Ballance, Karen Haskell, John P Noe, Dennis C Dinge, Thomas A Gardiner, and Michael E Davis. 2008. *7X Performance Results–Final Report: ASCI Red vs. Red Storm*. Technical Report. Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States).
[11] Ranga Raju Vatsavai and Varun Chandola. 2016. Guest editorial: big spatial data. *GeoInformatica* 20, 4 (2016), 797–799. https://doi.org/10.1007/s10707-016-0269-7
[12] Ranga Raju Vatsavai, Auroop R. Ganguly, Varun Chandola, Anthony Stefanidis, Scott Klasky, and Shashi Shekhar. 2012. Spatiotemporal data mining in the era of big spatial data: algorithms and applications. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, BigSpatial@SIGSPATIAL 2012, Redondo Beach, CA, USA, November 6, 2012*. 1–10.
[13] Yuanfeng Wu, Lianru Gao, Bing Zhang, Bin Yang, and Zhengchao Chen. 2015. Embedded GPU implementation of anomaly detection for hyperspectral images. In *SPIE Remote Sensing*. International Society for Optics and Photonics, 964608–964608.