

Unlocking Value from Ubiquitous Data

Rajendra Akerkar and Minsung Hong

Western Norway Research Institute,
P.O.Box 163, NO-6851 Sogndal, Norway
{rak, msh}@vestforsk.no

Abstract. Data is growing at an alarming rate. This growth is spurred by varied array of sources, such as embedded sensors, social media sites, video cameras, the quantified-self and the internet-of-things. This is changing our reliance on data for making decisions, or data analytics, from being mostly carried out by an individual and in limited settings to taking place while on-the-move and in the field of action. Unlocking value from data directs that it must be assessed from multiple dimensions. Data's value can be primarily classified as information, knowledge or wisdom. Data analytics addresses such matters as what and why, as well as what will and what should be done. In recent days, data analytics is moving from being reserved for domain experts to becoming necessary for the end-user. However, data availability is both a pertinent issue and a great opportunity for global businesses. This paper presents recent examples from work in our research team on ubiquitous data analytics and open up to a discussion on key questions relating methodologies, tools and frameworks to improve ubiquitous data team effectiveness as well as the potential goals for a ubiquitous data process methodology. Finally, we give an outlook on the future of data analytics, suggesting a few research topics, applications, opportunities and challenges. This paper is based on a keynote speech to the 14th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer, Kyiv, Ukraine on 16 May 2018.

Keywords: Ubiquitous data, Big data, Data analytics, Transport sector, Smart city, Emergency management

1 Introduction

Data are limitless and inexhaustible can be re-used multiple times for different uses. Supply of data is growing exponentially and delivers far more riches and value (knowledge) than any other assets. Data growth at an alarming rate is spurred by a variety of sources, such as embedded sensors, social media sites, video cameras, the quantified-self and the internet-of-things [1]. This is changing our reliance on data for decision making or data analytics, from being mostly carried out by an individual and in limited settings, to taking place while on-the-move and in the field of action. In terms of data evolution, data converges into wisdom (or intelligence) by way of information and knowledge through activities

like researching, absorbing, acting, interacting and reflecting [2]. While conducting these activities, persons as well as businesses generally gains understanding, experience and insight, and may come up with innovative ideas.

On the one hand, organisations historically collected copious amounts of data. However, by complexity of data analytics, the organisations were unable to use that data to generate meaningful information in a timely manner, and their business insights were fragmented [3]. According to Figure 1 that shows a value chain as the foundation of Big Data, these situations mean that there are difficulties predicting and responding to changing business needs and rising chances, as a result, business opportunities and related growth were tied to a much slower roadmap. Some Big Data experts have argued that businesses do

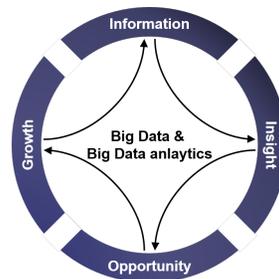


Fig. 1. Value chain in Big Data

not use even a small portion of Big Data that they have, and big does not always mean better [4, 5]. In other words, as one of data availability aspects, having tremendous amounts of data might not guarantee businesses a competitive advantage over competitors. Therefore, how effectively and quickly you analyse the data and extract actionable information from it will be critical, since fast and actionable data will become an important part in the usage of Big Data [6, 7].

In this regard, new data-and-analytics-related businesses and the application of data insights are changing the nature of competition. Data and analytics are also changing the nature of industry competition. Recently, McKinsey reports¹ that seventy percent of all executives think that Big Data and its analytics have caused changes in their industries competitive landscapes in recent years. In The fourth industrial revolution, Big Data analytics can support a productivity leap because it generates so many distinct opportunities². One of the simplest forecasts is faster change pace, changes happen more quickly and organizations can do more things in less time. Additionally, higher efficiency means that these require fewer resources, while enhanced effectiveness offers the changes greater effect. For instance, Increased predictability, through more accurate forecasting

¹ McKinsey&Business: Retrieved from <https://www.mckinsey.com/business-functions/operations/our-insights/ops-4-0-fueling-the-next-20-percent-productivity-rise-with-digital-analytics>

² McKinsey&Business: Retrieved from <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/analytics-comes-of-age>

based on unstructured data, lets organizations plan their moves more consistently and respond with greater agility. There are tremendous opportunities for organizations to leverage data in new ways yet many challenges still exist as follows: Many organizations fail to identify the right data or have no idea how to best use them. In addition, insufficient storage capacity and analytical capabilities to handle the massive volume of data. Furthermore, many enterprises lack the right kind of infrastructure or connectivity that can provide seamless access to data. Since data may come in a variety of formats, organizations often find it difficult and expensive to analyse and dig up insights. Last data is generated at the exponential rate of velocity, which makes it difficult to perform real-time analytics that uncovers the intelligence they contain.

In terms of “variety” in characteristics of Big Data, it’s analytics can be applied to many different types of data. These data types range from simple atomic types to more complex constructs such as time series, relational data, graphs, images and many others. Over the years, many methods were studied to transform into information and incorporate these different data types to get knowledge. One aspect that has been studied only briefly so far is ubiquity. In this regard, there are following characteristics [8].

- Data is producing asynchronously in a highly decentralized way.
- Data usually is involved to many different data types that make up a whole.
- Data in almost all cases emerges from a very high number of partially overlapping and loosely connected sources.
- Data is produced by many different users distributed all over the world and is often noisy and contradicting, partially overlapping etc.
- Data is multi-dimensional and multi-modal.

We refer to such data as “*Ubiquitous Data*” and define by:

the data emerges from pervasive domain specific information provided by static and mobile professional sensors, from social sensors -people- who voluntarily provide information/data, from large-scale data banks of business, municipalities and government, from social media, and open data.

This paper presents recent examples from work in our research team on LeMO³ and BDEM⁴ projects related to ubiquitous data analytics and open up to a discussion on key questions relating sources, methodologies, tools and frameworks to improve ubiquitous data team effectiveness as well as the potential goals for a ubiquitous data process methodology. Finally, we give an outlook on the future of data analytics, suggesting a few research topics, applications, opportunities and challenges.

³ LeMO: Retrieved from <https://lemo-h2020.eu/>

⁴ BDEM: Retrieved from <https://bdem.squarespace.com/>

2 Related Work

This section shows recent examples such as projects, methodologies, tools, frameworks, related to Ubiquitous Data. MK:smart⁵ as a large collaborative initiative has created a ‘MK Data Hub’ which supports the acquisition and management of vast amounts of data relevant to the city of Milton Keynes, England, systems from a variety of data sources. These have 712 data sets including local/national open data and data streams from both key infrastructure networks (energy, transport, water) and other relevant sensor networks (e.g. weather and pollution data), crowd-source data from social media and mobile applications. It provides a streaming API for timeline and sensor data, as well as an Entity API which aggregates the information available in the hub on general entities from the many data sets contributing to the MK Data Hub by using semantic technology. Especially the Entity API might offer a way of fusing various data sets to generate information from ubiquitous data toward knowledge. UBIMOB⁶ develops an adaptive and context sensitive mobility solution in order to make smart decisions taking citizens’ personal need into account and to reach equilibrium of mobility services, supply and demand, by smarter resource planning and matchmaking. Aiming to enable a vision for future mobility, this project considers huge amounts of ubiquitous data through sensors, traffic cameras, as well as asynchronous user-generated information, synchronous user-generated data, historic databases and data from mobility companies in real-time. Their solution will be evaluated by people, business and decision makers in 3 cities in Norway. SETA⁷ is developing technologies and methodologies based on large, complex and dynamic ubiquitous data from citizens, connected cars, city sensors and distributed databases, in order to change the way mobility is organised, monitored and planned in large metropolitan areas. Their application using smartphones sensors enable to track users’ mobility with considering you an insight into your daily travel. In addition, by integrating with environmental data such as road closures, road works, accidents, delays and levels of pollution, it shows users’ daily journeys, the impact they had on the environment, calories burned and any costs associated with the journey. TransformingTransport⁸ project is trying to demonstrate measurable, and replicable way the transformations that Big Data will bring to the mobility and logistics market. It addresses seven major domains, such as high-ways, sustainable vehicle, proactive rail infrastructures, intelligent ports, efficient air transport, multi-modal Urban mobility, and dynamic supply chains. Moreover, an open data portal⁹ containing 148 data sets related to ubiquitous data has been published in order to provide the community for the reuse of data across the different transport domains. NOESIS¹⁰ aims to offer a methodological

⁵ All information can be accessed via <http://www.mksmart.org/> and were last accessed on September 3, 2018.

⁶ <https://www.vestforsk.no/en/project/ubiquitous-data-driven-urban-mobility>

⁷ <http://setamobility.weebly.com/>

⁸ <https://transformingtransport.eu/>

⁹ <https://data.transformingtransport.eu/> was accessed on September 3, 2018

¹⁰ <http://noesis-project.eu/>

framework (i.e. a decision support tool) and data-driven evidence to enable the deployment of a Big Data in transport ecosystem in Europe, by addressing the related technological, institutional/legal, business, and policy challenges. Based on a number of machine learning techniques to associate Ubiquitous Data with a predefined set of use cases, the framework/tool will be used to assess the value generated (i.e. socioeconomic impact) from Big Data investments. In order to enhance the economic sustainability and competitiveness of European transport sector, LeMO¹¹ studies and analyses Big Data in the European transport domain in particular with respect to five transport dimensions: mode, sector, technology, policy and evaluation. Contrary to the NOESIS project that exploring case studies after implementing a framework, LeMO conducts a series of case studies to provide recommendations and roadmap on the prerequisites of effective Big Data implementation in the transport field, towards data openness, collection, exploitation and data sharing to support European transport stakeholders. In addition, this project considers economic, legal, ethical, political and social issues related to Ubiquitous Data in the field of transport. BDEM¹² aims to develop a international partnership between four countries (i.e. Norway, the USA, Japan and Hong Kong) to share best practices, and strengthen research in Big Data and emergency management. Another major initiative in our research group is Teknoløft project (research-based technology innovation in Sogn og Fjordane region) that aims to nurture and develop a strong knowledge platform on Big Data, which enables local business to be able to use existing and new data for innovation.

So far, some projects related to Ubiquitous Data (especially smart city, smart transport and emergency management) are introduced. In common, the projects argue that Ubiquitous Data should be available, accessible and reusable in order to create more invaluable information and knowledge. In addition, various issues (e.g. economic, legal, social, ethic, environmental issues) in target field also should be considered to appropriately exploit Ubiquitous Data. More detail challenges and open issues will be discussed in Section 4.

3 Understanding Ubiquitous Data

The advent of Ubiquitous Data is explained by two parallel movements. One is technological change, for instance cloud computing enables large volumes of data to be stocked at low cost. The improved speed of data update and analysis (real time) permits significant optimisations in all most economic sectors such as transport, energy, construction, health and so on. In addition the advance of many techniques in a variety of fields (e.g. machine learning, data mining and artificial intelligence) are allowing to deal with various data forms and formats likes images, texts, sounds and so on. Other change is in behaviour and use. Nowadays, by the increase in Internet and social media usage, data produced every day on the web represents the equivalent of information contained on 250

¹¹ <https://lemo-h2020.eu/overview/>

¹² <https://bdem.squarespace.com/>

million DVDs; On the horizon in 2020, there should be 50 million connected devices globally (smart phones, tablets, smart watches and etc.) These movements for Ubiquitous Data are addressing potential to add value as follows:

1. Data sharing between stakeholders at different levels, but whose interests converge towards certain projects.
2. Crossing of different types and forms of data to keep creating more reliable and easier-to-use services.
3. Interoperability of data, that is the technical capability to connect data of different types and origins.
4. Citizen participation according to their right for the data that they produce and want to put to good use.

These capabilities may be obtained, when following principles [9] of Ubiquitous Data are sufficiently figured out.

- **Data comes from activities.** For instances, every movement of a freight, change of factory temperature, or adjustment of a wing flap is an activity. In most of the machines and buildings of the world, data/information created without these activities is easily lost. Companies are in a race to digitize and datafy them before their rivals do. Digitizing activities means involving sensors or mobile applications in the activities in some way, and datafying activities means extending the observations you capture about them.
- **Data tends to make more data.** As one of examples, algorithms that use accumulated data to predict maintenance schedules for electrical transformers or inventory movements in automated warehouses produce data about their own performance in order to be fed back into the system and improve their future performance. This may becomes a competitive lead that's very hard for others to close. Businesses must develop both infrastructure and new skills in experiment design, data analysis, and interpretation. Managing that process will require new tactics to handle conflicting, data-driven perspectives on the same topic.
- **Platforms being preferred tend to win.** The digitization and datafication of more mechanized activities brings platform competition to industries that have not seen it before. Already, luxury car makers are in a battle to be the preferred platform for connected vehicle services. This same phenomenon will change the terms of competition for wide areas from manufacturers to construction.

Following sections present sources and processing of Ubiquitous Data in order to glance these potentialities.

3.1 Ubiquitous Data Sources

3.1.1 Mobility Data Among mobility data, mobile data rose in prominence with the advance of smart phones and the interest in civic apps. A call record data containing attributes that are specific to a single instance of a phone call

(or data connection) such as the phone numbers and the start/end time and the duration of that call, as well as sensor data which generated embedded sensors to smartphone, even applications installed in users mobile have been utilised in various fields (e.g., analysing individuals' behaviour [10], building a network between different users [11] and inferring other patterns [12]). For instance, this kind of data can be studied to understand large-scale human mobility and trajectory patterns [13]. As another mobility data, data generated by moving vehicle traveling around a city with a GPS sensor is sent to a central/distributed system and matched to a road/rail/path network for deriving and managing traffic status. As becoming this kind of data vary and complex by the emergence of (semi-)automated vehicles and drones, its utilisation is being studied widely for various areas such as transport optimisation [14], emergency management [15], sustainable city [16]. Commuting/public transport data, such as the card usage data for systems of a bus and subway and the ticketing data in parking lots, is generated by people living in cities. This kind of data generated by people passing entrances of a subway station and getting a bus can be used for public transportation systems in a city [17]. Each transaction record contains timestamps and the ID of the station or bus as well as the fare for this move. Payment information created from parking meters in street-side may include the time the parking ticket and the fare. These data indicate the traffic of around a parking place, which can be used to not only manage a city's parking infrastructure [18] but also analyse people's travel patterns [19].

3.1.2 Environmental Data Meteorological data enable to be crawled from public websites includes various environmental information such as humidity, temperature, barometer pressure, wind speed, and weather conditions [20]. It influences traffic flow and land uses. On the one hand, there are many ways to collect traffic data, such as using loop detectors, magnetometers, infrared sensors, surveillance cameras and so on [21, 22]. For example, the loop detectors are typically buried in the major roads (e.g. highways). Such detectors deduces the travel speed of vehicles by using two consecutive (i.e. a pair of) sensors. In addition, the traffic volume on a road can be detected by counting the number of vehicles traversing the loop detectors. However, its cost is usually expensive and area in which the sensors are embedded and limited [23]. Therefore other data can be gathered via infrared or ultrasonic sensors. Additionally, some traffic surveys can be conducted manually [24]. As another instance, surveillance cameras, installed in cities over the last thirty or so years, generate a huge amount of image and video data reflecting traffic patterns. However, it is still challenge to automatically extract traffic information from the collected data using machine learning techniques [25]. Hence, monitoring citywide traffic status is mainly conducted through human resource.

3.1.3 Social Media Data Due to the advent of the personal mobile devices and web 2.0, Social media data consisting of social network (e.g. relationship, interdependency, or interaction) and user-generated data (e.g., texts, photos and

videos) has being utilised in the wide fields such as transport sector [26], smart city [27] and emergency management [28]. It includes enriched information for ordinary or specific context and users' behaviour/interests. In particular, some part of the data have user's location and timestamp. Therefore many researchers and practitioners have been tried to use it for their domains. However, there still are many issues such as reliability, processing unstructured data [29], sparsity of locations [30] and privacy. For instance, it is difficult to obtain reliable data from social media, since it is usually many noisy and contain too broad information from daily life to specific context [31]. In addition, social media, since it is content directly created by end users, might carry personally identifiable information. Therefore, researchers and practitioners should be mindful of any explicit or inferred personal information in the data [32].

3.1.4 Geographical Data Road network data is one of the most frequently utilised geographical data in various area, for instatnce, transport routing [14], city planning [33] and emergency management [34]. This kind of data is able to be presented by a graph consisting of edges (i.e. road segments) and a nodes (intersections). An edge explaining by two nodes and a geospatial points is related to the length, speed constraint, tyep of road and number of lanes, while a node has geospatial coordinates. On the on hand, a Point of interest (POI), such as a restaurant, a museum and a zoo, usually has a name, address, category and geospatial coordinates. It is difficult to gather POI data due to a characteristic of their changing according to time. For instance, a shop may replace its name, be moved to a new place, or even shut down. The POI data can be generated by thwo methods: one is collected from existing Yellow page data by geocoding algorithm, and the other is manually gathering POI information (e.g., Navinfo and AutoNavi). In case of Foursquare, as a location-based social networking service, end users is able to create a new POI which has not been included in a system. Online map services (e.g., Google maps and Bing) generally combine the aforementioned two approaches to collect a large coverage of POIs. Therefore, issues have been arisen, such as accuracy of POI [35], integrating the POI data from various [36] sources and so on.

3.2 Processing Ubiquitous Data

3.2.1 Architecture for Ubiquitous Data: In this subsection, we propose an architecture of Ubiquitous Data Analytics platform as shown in Figure ?? and briefly describe each component of the architecture. It consists of four layers such as data collection, analytics, management and providing.

- **Data collection layer** gathering data is connected closely to physical world and can often be called as edge computing (or network). Heterogeneous data are generated from a variety of sources and transferred to edge clouds (or routers) in order to conduct learning algorithms in real-time. Online or ensemble learning technique has being studied as an appropriate approach

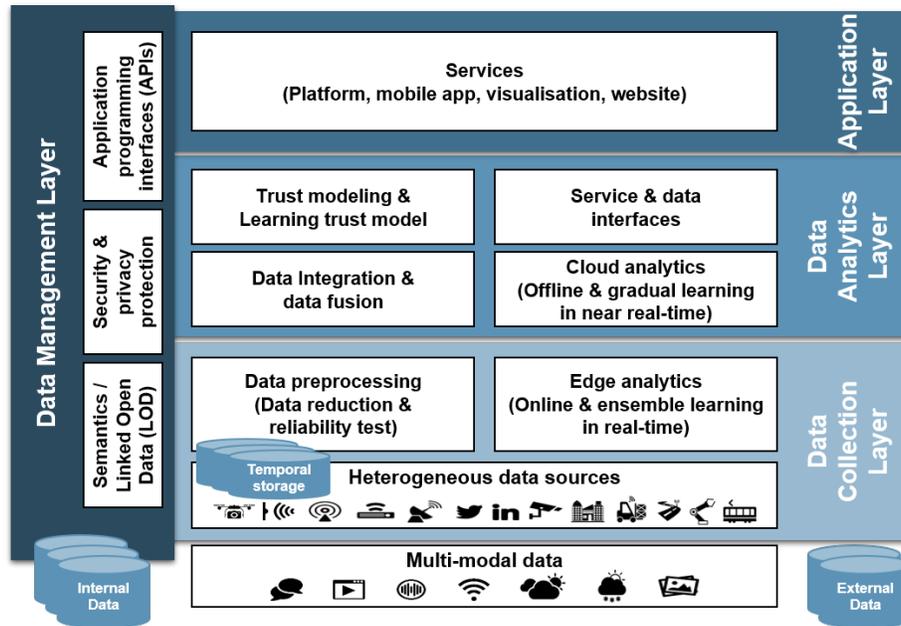


Fig. 2. An architecture for Ubiquitous Data Analytics platform

for edge analytics. Data preprocessing such as data reduction and reliability test is necessary, since edge nodes have limited resources.

- **Data analytic layer** obtain various types of data from edge nodes to analyse data using powerful resources (i.e., storage, computing power and so on) in near real-time. This layer also performs data integration and fusion through semantic and LOD techniques in order to obtain higher level of information and knowledge. In addition, trust modelling and its learning are conducted to a many amount of data based on the strong resources. The model learnt may be propagated at each edge to test data reliability in real time. Last refined or learnt data/information will be presented through service and data interfaces to the application layer (e.g., platform, mobile app, visualisation, website).
- **Application layer** directly connected with service providers or end users to offer data/information. Here visualisation techniques opening a way of efficient and effective information provision have being studied and utilised.
- **Data management layer** impacts other layers overall. Aforementioned semantic/LOD techniques can also be used in data collection and data providing layers. For instance, a visualisation generating data network may utilise LOD technique to discover relations (i.e. edges of graph) between data nodes. Another important issue is preserving data security & privacy and must be performed for all process carefully. Application programming

interfaces should be also offered to enhance the accessibility and usability of Ubiquitous Data.

Next subsection presents more details for data management and analytics.

3.2.2 Data Management: Ubiquitous Data management involves development and execution of architectures, policies, practices and procedures that properly manage the full data lifecycle needs for applications using Ubiquitous Data [37]. As data comes from different sources with different formats, there is a need for advanced data management features that will lead to recognizing the different formats and sources of data, structuring, managing, classifying and controlling all these types and structures. Ubiquitous data management should also provide scalable handling for massive data to support offline applications as well as low latency processing to serve effectively in real-time applications. In addition, it is a policy-based approach for determining which information should be stored where within an organization's IT environment, as well as when data can safely be deleted. Within a typical enterprise, people with many different job titles may be involved in ubiquitous data management. They are a chief data officer, chief information officer, data managers, database administrators, data architects, data modelers, data scientists, data warehouse managers, data warehouse analysts, business analysts, developers and others.

3.2.3 Data Analytics: Typical analytic algorithms used in regular applications may not be sufficient or efficient enough to handle Ubiquitous Data applications due to their unique requirements and pressing need for high volume high speed processing. These algorithms need to be optimized to handle high data volumes, large variety of data types, time constraints on decision making processes, and distributed components across various geographical locations. Also, these algorithms need to work effectively across heterogeneous environments and be capable of managing and operating in highly dynamic environments. Additionally, they should consider handling high data volumes, working across large variety of data types, time constraints on decision making processes, geographical distributed components [16]. Unobtrusive and ubiquitous sensing technologies, advanced data management and analytics models, and novel visualization methods, should be connected to create win-win-win solutions that improve urban environment, human life quality, and city operation systems.

4 Challenges and Open Issues

This section explores challenges and open issues on the Ubiquitous Data and its utilisation. As aforementioned, Ubiquitous Data is dynamically changed with respect to time and space and needs to be processed by considering its heterogeneous from various type of sources in real time. Actually, for instance, monitoring traffic flow on a road segment is possible, but continually probing

the citywide traffic is challenging as we do not have sensors on every road segment; Additionally, there is limited integration of ubiquitous data analytics and social computing; And there is an absence of effective models to predict behaviours efficiently; Real-time prediction and mobility capabilities have not yet been incorporated into the infrastructure for smart transport management. In this regard, following challenges should be considered in order to unlock real values (range from information to wisdom) of Ubiquitous Data.

Accessibility and usability of Open data: The open data movement is based on the three principles of transparency, participation and collaboration [38]. While it is believed that through openness, sharing and cooperating, the value of data to society can be truly realized, the rapid opening of government data has not been universally welcomed. However, there are still many open data portals simply providing dump files. These sites do not follow any standard or even give any attention for the usability of the data. One of the reasons is that they have been developed by hackathon enthusiasts or data divers [39]. In these cases, unfortunately, there is often no post-event follow up, maintenance or further development [40]. Openness usually refers a number of aspects that circle around the fact that an element is openly (in the sense of publicly and royalty free) available and reusable, developed in an open process, accessible at minimum costs (in terms of data pure reproduction costs or even no costs).

Processing and integration of heterogeneous data sources: Overlapping the same data, increasing performance and scalability and enabling real-time data access are one of the challenges related to data processing and integration that should be addressed in the future. For instance, collecting data from different sources, storing and providing data with a unified view, and adjusting structures in semi-structured and unstructured data [41]. Traditional data mining usually deals with data from a single domain. In the ubiquitous data era, we face a diversity of datasets from different sources in different domains. These datasets consist of multiple modalities, each of which has a different representation, distribution, scale, and density. How to unlock the power of knowledge from multiple disparate (but potentially connected) datasets is paramount in ubiquitous data research, essentially distinguishing ubiquitous data from traditional data mining tasks. This calls for advanced techniques that can fuse the knowledge from various datasets organically in a machine learning and data mining task.

Spatio-temporal property: Spatio-temporal locality is central feature as well as challenge of Ubiquitous Data. Temporality means whether the data describe phenomena that change with time, while spatiality indicates whether the data are spatially located. All most of Ubiquitous data have characteristics combined of both features. In a typical scenario related to Ubiquitous Data, learning algorithm has access to past and maybe present data from a time-varying distribution, and has to make inferences about the present or future. Additionally, it is assumed that a data stream is constantly generated with a memory constrained environment [42]. Thus, various techniques such as concept drift detectors, data reduction, sliding windows and online/ensemble learners have being studied [43].

In this regard, ‘*edge analytics*’ is relatively new and it is still developing [44]. Once it is sufficiently perfected, it may revolutionize the way we process ubiquitous data. Basically, data is analysed the moment it is collected, so you immediately have a complete analysis. This can be really useful for security cameras so that irrelevant data is no longer stored, for navigation devices and so on.

Privacy and security: A wide range of smart mobility technologies are being deployed within ubiquitous environment. These technologies are generating huge quantities of data and much of them in real-time. However, generating, processing, analysing, sharing, and storing vast amounts of actionable data also raises several concerns and challenges. For example, data privacy, data protection, and data security issues are caused from the creation of smart mobility [45]. Privacy-related issues arise when a system infers or restores personal information using Big Data analytics tools, although data are anonymized. With the proliferation of Big Data analytics technologies used in Ubiquitous data, the privacy issue has become a core problem in the data mining domain. Another security risk associated with Ubiquitous Data is the heterogeneity of the types of devices used and the nature of generated data (e.g., raw devices, data types and communication protocols). To authenticate these devices, a system should assign and maintain a nonrepudiable identification to each device. These activities results in increased security risks [41].

5 Conclusion

As growing Ubiquitous Data at an alarming rate by varied array of sources, such as embedded sensors, social media sites and the internet-of-things, this is changing our reliance on data for making decisions. Nowadays, data analytics is moving from being conducted by/for domain experts, to becoming necessary for the end-user. However, there are still many challenges such as data availability, processing heterogeneous data and privacy. In this paper, we presented recent examples (e.g. projects, methodologies, tools and frameworks) on Ubiquitous Data analytics and discussed on key question relating Ubiquitous Data to give an outlook on the future of data analytics, a few research topics, challenges and opportunities. In this regard, an architecture of Ubiquitous Data Analytics platform was also proposed by considering recent environments and research trends.

Acknowledgements

This work is partially supported by the European Unions Horizon 2020 research and innovation programme under grant agreement No 770038, UBIMOB project (270785) funded by the Norwegian Research Council in 2017 through the IK-TPLUSS programme, and BDEM project funded by the Research Council of Norway (RCN) and the Norwegian Centre for International Cooperation in Education (SiU) through the INTPART programme. Authors contributed equally to this work.

References

- [1] Akerkar R (2019) Processing big data for emergency management. In: *Emergency and Disaster Management: Concepts, Methodologies, Tools, and Applications*, IGI Global, pp 980–1000
- [2] Akerkar R, Sajja P (2010) *Knowledge-based systems*. Jones & Bartlett Publishers
- [3] Akerkar R, Sajja PS (2016) *Intelligent Techniques for Data Science*. Springer, DOI 10.1007/978-3-319-29206-9
- [4] Goes JD (2013) Big data is dead. whats next. Venturebeat-com guest blog post URL https://venturebeat.com/2013/02/22/big-data-is-dead-whats-next/?goback=%2Egde_62438_member_217099766
- [5] Chauhan R (2015) Transforming big data into actionable insights URL https://www.mastercardadvisors.com/content/dam/advisors/en-us/documents/150513_Transforming_Big_Data.pdf
- [6] Barnaghi PM, Sheth AP, Henson CA (2013) From data to actionable knowledge: Big data challenges in the web of things. *IEEE Intelligent Systems* 28(6):6–11, DOI 10.1109/MIS.2013.142
- [7] Carter KB (2014) *Actionable Intelligence: A Guide to Delivering Business Results with Big Data Fast!* John Wiley & Sons
- [8] Hotho A, Pedersen RU, Wurst M (2010) Ubiquitous data. In: May M, Saitta L (eds) *Ubiquitous Knowledge Discovery - Challenges, Techniques, Applications*, Lecture Notes in Computer Science, vol 6202, Springer, pp 61–74, DOI 10.1007/978-3-642-16392-0_4
- [9] Insights MTR (2016) The rise of data capital. Oracle, URL <https://www.technologyreview.com/s/601081/the-rise-of-data-capital/>
- [10] Senaratne H, Müller M, Behrisch M, Lalanne F, Bustos-Jiménez J, Schneidewind J, Keim DA, Schreck T (2018) Urban mobility analysis with mobile network data: A visual analytics approach. *IEEE Trans Intelligent Transportation Systems* 19(5):1537–1546, DOI 10.1109/TITS.2017.2727281
- [11] Song Y, Hu Z, Leng X, Tian H, Yang K, Ke X (2015) Friendship influence on mobile behavior of location based social network users. *Journal of Communications and Networks* 17(2):126–132, DOI 10.1109/JCN.2015.000026
- [12] Xia D, Lu X, Li H, Wang W, Li Y, Zhang Z (2018) A mapreduce-based parallel frequent pattern growth algorithm for spatiotemporal association analysis of mobile trajectory big data. *Complexity* 2018:2818,251:1–2818,251:16, DOI 10.1155/2018/2818251
- [13] Bhattacharya S, Blunck H, Kjærgaard MB, Nurmi P (2015) Robust and energy-efficient trajectory tracking for mobile devices. *IEEE Trans Mob Comput* 14(2):430–443, DOI 10.1109/TMC.2014.2318712
- [14] Menouar H, Güvenç I, Akkaya K, Uluagac AS, Kadri A, Tuncer A (2017) Uav-enabled intelligent transportation systems for the smart city: Applications and challenges. *IEEE Communications Magazine* 55(3):22–28, DOI 10.1109/MCOM.2017.1600238CM

- [15] Chen L, Englund C (2018) Every second counts: Integrating edge computing and service oriented architecture for automatic emergency management. *Journal of Advanced Transportation* 2018
- [16] Nuaimi EA, Neyadi HA, Mohamed N, Al-Jaroodi J (2015) Applications of big data to smart cities. *J Internet Services and Applications* 6(1):25:1–25:15, DOI 10.1186/s13174-015-0041-5
- [17] Ma X, Liu C, Wen H, Wang Y, Wu YJ (2017) Understanding commuting patterns using transit smart card data. *Journal of Transport Geography* 58:135–145
- [18] Bagula AB, Castelli L, Zennaro M (2015) On the design of smart parking networks in the smart cities: An optimal sensor placement model. *Sensors* 15(7):15,443–15,467, DOI 10.3390/s150715443
- [19] Zhao Z, Koutsopoulos HN, Zhao J (2018) Detecting pattern changes in individual travel behavior: A bayesian approach. *Transportation Research Part B: Methodological* 112:73–88
- [20] Alam F, Mehmood R, Katib I, Albogami NN, Albeshri A (2017) Data fusion and iot for smart ubiquitous environments: A survey. *IEEE Access* 5:9533–9554, DOI 10.1109/ACCESS.2017.2697839
- [21] Nandury SV, Begum BA (2015) Smart wsn-based ubiquitous architecture for smart cities. In: Mauri JL, Thampi SM, Wozniak M, Marques O, Krishnaswamy D, Sahni S, Callegari C, Takagi H, Bojkovic ZS, M V, Prasad NR, Calero JMA, Rodrigues J, Que X, Meghanathan N, Sandhu R, Au E (eds) *Proceedings of the International Conference on Advances in Computing, Communications and Informatics, ICACCI 2015*, IEEE, Kochi, India, pp 2366–2373, DOI 10.1109/ICACCI.2015.7275972
- [22] Yi S, Li C, Li Q (2015) A survey of fog computing: Concepts, applications and issues. In: Li Q, Xuan D (eds) *Proceedings of the 2015 Workshop on Mobile Big Data, Mobidata@MobiHoc 2015*, ACM, Hangzhou, China, pp 37–42, DOI 10.1145/2757384.2757397
- [23] Thakuriah PV, Geers DG (2013) Data sources and management. In: *Transportation and Information*, Springer, pp 15–34
- [24] Taylor N, Stott I, Parker J, Bradley J, Graham A, Tuppen C, Morley J (2015) The transport data revolution: Investigation into the data required to support and drive intelligent mobility URL <https://ts.catapult.org.uk/wp-content/uploads/2016/04/The-Transport-Data-Revolution.pdf>
- [25] Chen N, Chen Y, You Y, Ling H, Liang P, Zimmermann R (2016) Dynamic urban surveillance video stream processing using fog computing. In: *IEEE Second International Conference on Multimedia Big Data, BigMM 2016*, Taipei, Taiwan, April 20-22, 2016, IEEE Computer Society, pp 105–112, DOI 10.1109/BigMM.2016.53
- [26] Anantharam P, Barnaghi PM, Thirunarayan K, Sheth AP (2015) Extracting city traffic events from social streams. *ACM Transactions on Intelligent Systems and Technology* 6(4):43:1–43:27, DOI 10.1145/2717317
- [27] Costa DG, Duran-Faundez C, Andrade DC, Rocha-Junior JB, Peixoto JPJ (2018) Twittersensing: An event-based approach for wireless sensor net-

- works optimization exploiting social media in smart city applications. *Sensors* 18(4):1080, DOI 10.3390/s18041080
- [28] Poblet M, García-Cuesta E, Casanovas P (2017) Crowdsourcing roles, methods and tools for data-intensive disaster management. *Information Systems Frontiers* pp 1–17
- [29] Luna S, Pennock MJ (2018) Social media applications and emergency management: a literature review and research agenda. *International journal of disaster risk reduction* 28:565–577
- [30] Burton SH, Tanner KW, Giraud-Carrier CG, West JH, Barnes MD (2012) ” right time, right place” health communication on twitter: value and accuracy of location information. *Journal of medical Internet research* 14(6)
- [31] Kim J, Hastak M (2018) Social network analysis: Characteristics of online social networks after a disaster. *Int J Information Management* 38(1):86–96, DOI 10.1016/j.ijinfomgt.2017.08.003
- [32] Imran M, Castillo C, Diaz F, Vieweg S (2015) Processing social media messages in mass emergency: A survey. *ACM Computing Surveys (CSUR)* 47(4):67
- [33] Yin C, Xiong Z, Chen H, Wang J, Cooper D, David B (2015) A literature survey on smart cities. *SCIENCE CHINA Information Sciences* 58(10):1–18, DOI 10.1007/s11432-015-5397-4
- [34] Yin J, Yu D, Yin Z, Liu M, He Q (2016) Evaluating the impact and risk of pluvial flash flood on intra-urban road network: A case study in the city center of shanghai, china. *Journal of hydrology* 537:138–145
- [35] Ko EB, Lee JW (2014) Accuracy improvement methods for string similarity measurement in poi (point of interest) data retrieval. *KIISE Transactions on Computing Practices* 20(9):498–506
- [36] Jiang S, Alves AO, Rodrigues F, Jr JF, Pereira FC (2015) Mining point-of-interest data from social networks for urban land use classification and disaggregation. *Computers, Environment and Urban Systems* 53:36–46, DOI 10.1016/j.compenvurbsys.2014.12.001
- [37] Mosley M, Brackett MH, Earley S, Henderson D (2010) DAMA guide to the data management body of knowledge. Technics Publications
- [38] Lathrop D, Ruma L (2011) Open government: Collaboration, transparency, and participation in practice. *Government Information Quarterly* 28(1):129–130, DOI 10.1016/j.giq.2010.08.001
- [39] Townsend AM (2013) Smart cities: Big data, civic hackers, and the quest for a new utopia. WW Norton & Company
- [40] Barkham R, Bokhari S, Saiz A (2018) Urban big data: City management and real estate markets. *GovLab Digest*: New York, NY, USA
- [41] Marjani M, Nasaruddin F, Gani A, Karim A, Hashem IAT, Siddiqa A, Yaqoob I (2017) Big iot data analytics: Architecture, opportunities, and open research challenges. *IEEE Access* 5:5247–5261, DOI 10.1109/ACCESS.2017.2689040
- [42] May M, Berendt B, Cornue A, et al (2008) Research challenges in ubiquitous knowledge discovery. In: *Next Generation of Data Mining*, Chapman and Hall/CRC, pp 154–173

- [43] Ramírez-Gallego S, Krawczyk B, García S, Wozniak M, Herrera F (2017) A survey on data preprocessing for data stream mining: Current status and future directions. *Neurocomputing* 239:39–57, DOI 10.1016/j.neucom.2017.01.078
- [44] Satyanarayanan M, Simoens P, Xiao Y, Pillai P, Chen Z, Ha K, Hu W, Amos B (2015) Edge analytics in the internet of things. *IEEE Pervasive Computing* 14(2):24–31, DOI 10.1109/MPRV.2015.32
- [45] Akerkar R (2019) Privacy and security in data-driven urban mobility. In: *Utilizing Big Data Paradigms for Business Intelligence*, IGI Global, pp 106–128