

The two-way impact between data mining and mobile technologies

Mefail Tahiri

Docent at State University of Tetova ;
mefailt@gmail.com

Ejup Rustemi

Master at State University of Tetova:
ejuprustemi@yahoo.com

Abstract

Mobile devices have changed almost everything in the way how we communicate, work and deal with data. In some way they have facilitated businesses around the world, but in many ways they have produced a vast amount of data that most of the time are very complex and difficult to deal with. Data mining is a valuable tool in dealing with this issue. This paper will try to explain the relationship between data mining and mobile devices and technologies and the possibilities that they offer to improve the way we operate in small or large enterprises, schools, etc.

Key words: data mining, mobile technologies, mobile devices.

Introduction

Data mining has been for a long time one of the most important tools for analysing, retrieving, etc., large amounts of data from databases. It has provided and is still provides methods by which we can extract valuable information from our data. On the other hand, mobile devices are a relatively new technological trend, but in no way less important in the IT realm. Mobile technologies have had tremendous speed of development and nowadays it is impossible to imagine our world without it. "Data mining is a fundamental concept in the field of Information Retrieval. It is the process of analysing vast amounts of data in order to discover meaning beyond its explicit face value. For example, after collecting the browsing history of a user over a period of three months we could define a list of websites that the user visited and the time at which those sites were visited. A very simple application of data mining might be to then count the frequency with which each site was visited, and from this determine the user's favourite websites. We could also look at patterns in the time at which certain sites were visited; if the user visits the Sydney Morning Herald's website at 9am each morning this is a trend that we might want to incorporate into a user profile somehow. More complex patterns might also be present, allowing for work flows to be designed around a user's habits. These are just a few trivial examples."(1)

Mobile devices and the Internet

As IT has progressed, people have found new ways to communicate and do business with each other. The continuous need for faster and more effective results has paved the way for the introduction of mobile devices. Today humankind travels around the world as it travelled around a country in the past. Trying to be updated about how things progress has made mobile devices a very important tool, because they can be carried on ones pocket, you can access the Internet with them, read emails and reports. In one word, in our competitive society mobile devices keep us on the line with what happens around us.

“Mobile devices have opened up access to new, richer types of data that were once difficult, if not impossible, to obtain. Users of mobile devices can record their surroundings with visual and audio capture devices, transmit their location to online services with GPS-capable devices, and provide a myriad of other useful contextual cues for software services to use in the construction of a user profile. Dedicated devices were once required for each of these forms of data, but with the power of modern mobile phones it is becoming easier to generate this data with a single device and provide a more complete ubiquitous experience. Applying data mining to this rich data allows for the construction of complex user profiles, or for the delivery of sophisticated services.”(1)

The Internet and mobile devices cannot be separated from each other. They represent the main link that we have with the world.

“The internet and web services usage on mobile devices are continuously and rapidly increasing. Therefore the demand is to have efficient mobile interface that can effectively display information and efficiently utilize the small size mobile screen, low bandwidth and unreliable connection etc. Mobile devices, such as PDAs, Smartphones, Tablets etc. evolve rapidly from the digital calendars and address book to hosts of more complex functionalities. Because of improving mobile devices resources, it supports and adapting all forms of application from computer systems. However mobility and scaled-down technologies lead to some limitations as compare to computer system. Due to these limitations and deficiencies it is not straight forward to adopt or provide available facilities on mobile device as with desktop computers.”(2)

As we mentioned before all of this technological progress has given us an opportunity to do things better, but this has to do mainly with the end user, because there are people who need to do analysis of that big “tower” of data, or as IT experts call it, Big Data.

Software engineers have developed many applications that deal with the issue, but still, it requires a lot of work. Data analysis today is vastly superior than in the past, but there’s always something better than we can do.

“Analysis of data used for mobile is a complex process that often involves remote resources (computers, software, databases, files, etc.) and people (analysts, professionals, end users). Recently, mobile data mining techniques are used to extract useful data sets. Advancement in this research area arises from the use of mobile computing technology for supporting new data analysis techniques and new ways to discover knowledge from every place in which people operate. The availability of client programs on mobile devices that can invoke the remote execution of data mining tasks and show the mining results is a significant added value for nomadic users and organizations that need to perform analysis of data stored in repositories far away from the site where users are working, allowing them to generate knowledge regardless of their physical location. Aim of proposed work is to improve the mobile data mining techniques so that data retrieval for mobile devices will be faster in efficient mobility management using proper web services.”(3)

“We’ve witnessed several fast-moving revolutions in the technological landscape in the past few years, particularly related to the emergence of powerful, always-connected, and extremely popular portable devices. Smartphones and tablets let us receive information through multiple channels while generating

massive amounts of information about us. Data collected from the sensors embedded in smartphones — especially GPS receivers — provide an incredible wealth of information that service providers and applications can collect, store, and analyze in real time.”

It is personal

Today almost anyone can have access to information on their own on a cell phone, tablet, or even smartwatch.

“Computing is truly personal now, not only because we access information through mobile devices, but also because the information itself is usually highly personalized and relevant to our location and context. Typical examples are location-based services, such as Foursquare, which provide suggestions about restaurants and shops close to the area where users have checked in, and considers their previous mobility history. Other examples include search engines that are increasingly context- and location-aware. Moreover, users generate information themselves using mobile devices. For example, in June 2013, Facebook had, on average, 819 million monthly active mobile users. In other words, we should be talking not about the big data revolution but — at least as far as consumer applications are concerned — about the “big mobile data” phenomenon.”(4)

“The availability of client programs on mobile devices that can invoke the remote execution of data mining tasks and show the mining results is a significant added value for nomadic users and organizations that need to perform analysis of data stored in repositories far away from the site where users are working, allowing them to generate knowledge regardless of their physical location.”(5)

Domenico Talia and Paolo Trunfio have this to add in relation with data mining and mobile devices:

“The mobile data mining field may include several application scenarios in which a mobile device can play the role of data producer, data analyzer, client of remote data miners, or a combination of them. More specifically, we can envision three basic scenarios for mobile data mining:

- The mobile device is used as terminal for ubiquitous access to a remote server that provides some data mining services. In this scenario, the server analyzes data stored in a local or distributed database, and sends the results of the data mining task to the mobile device for its visualization.
- Data generated in a mobile context are gathered through a mobile device and sent in a stream to a remote server to be stored into a local database. Data can be periodically analyzed by using specific data mining algorithms and the results used for making decisions about a given purpose.

Mobile devices are used to perform data mining analysis. Due to the limited computing power and storage space of today’s mobile devices, currently it is not realistic to perform the whole data mining task on a small device. However, some steps of a data mining task (i.e., data selection and pre-processing) could be run on small devices.”(5)

Mobile invasion

The following table will give us a very good view of how large is the usage of mobile devices, and in return, we will understand the importance of data mining in dealing with data produced by all those numbers.

No.	Company	Country	Subscribers in million
1.	China Mobile	China	775.6
2.	Vodafone	UK	419.4
3.	China Unicom	China	285.7
4.	Airtel	India	275.2
5.	America Movil	Mexico	269.9
6.	Telefonica	Spain	254.7
7.	Axiata	Malaysia	239.7
8.	Orange	France	231.5
9.	VimpelCom Ltd.	Russia	209
10.	China Telecom	China	185
11.	MTN Group	South Africa	175.9
12.	Etisalat	UAE	167
13.	Telenor	Norway	166
14.	TeliaSonera	Sweden/Finland	160
15.	T-Mobile	Germany	142.5
16.	Saudi Telecom Company (STC)	Saudi Arabia	139
17.	Reliance Communications	India	135.8
18.	Verizon Wireless	USA	122
19.	Idea Cellular	India	113.9
20.	AT&T Mobility	USA	116
21.	MTS	Russia	106
22.	Telecom Italia Mobile (TIM)	Italy	102.5
23.	BSNL	India	96.2
24.	Tata Teleservices	India	77.4
25.	Smart Communications	Philippines	72.5
26.	Turkcell	Turkey	70.7
27.	Aircel	India	66.9
28.	Maxis Communications	Malaysia	63.7
29.	MegaFon	Russia	62
30.	Ooredoo	Qatar	60.5

Table 1. Mobile Services by the numbers (2013; data by RBI, Financial Access)(6)

These huge numbers paired with the continuous improvement of mobile devices (fig.1) are a good representation of the importance of data mining of/on mobile devices.

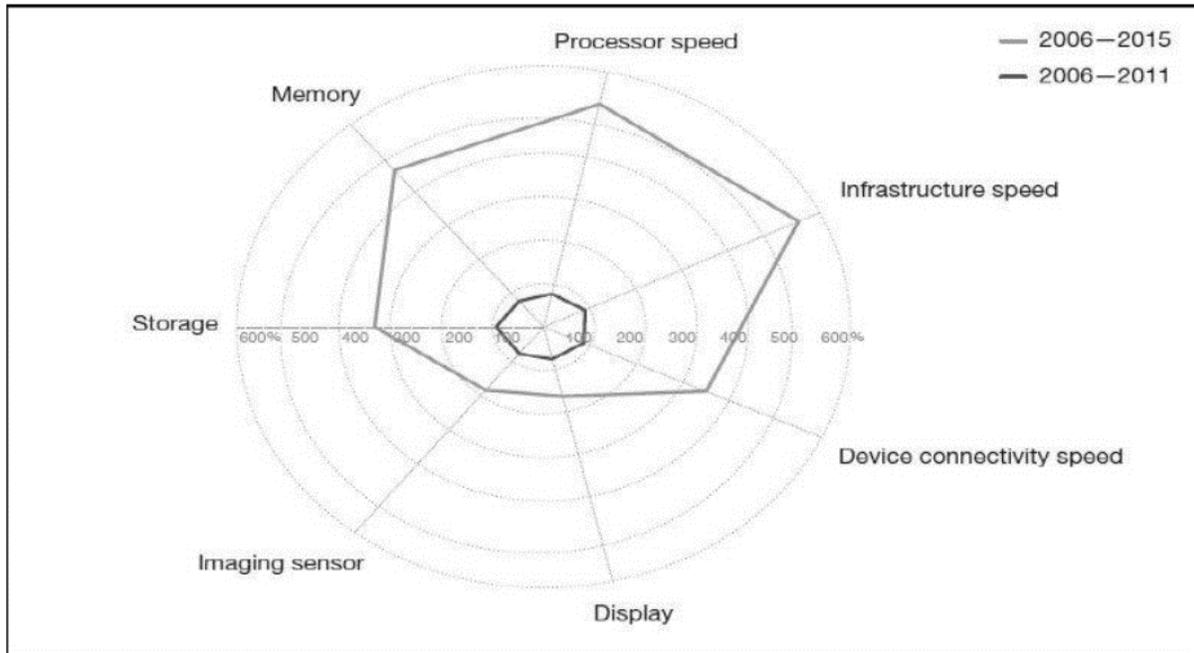


Fig.1 Development of mobile devices by year(6)

Is there a downside?

As with everything else, mobile devices have their downsides which in some cases can limit the ability to do things appropriately. This also effects data mining no matter if it is carried on the mobile devices themselves or if we try to mine mobile users using those devices. Here we will show some of the limitations as presented by Muzammil Khan, Ali Shah and Israr Ahmad on their 2014 paper "Framework for Interactive Data Mining Results Visualization on Mobile Devices".

"Despite of rapid development in mobile technologies, mobile devices have potential differences as compared to conventional desktop computer systems. Many of these differences are in the form of limitations. The number of limitations or constraints discussed in regarding mobile devices can be categorizing by the relevancy among these limitations. We can categorize into four different categories, that as Bandwidth or Network related constraints, Screen or Display related constraints, Hardware or Software constraints and Mobility or Location Related constraints."

Network Constraints

- Slow Connectivity
- Temporary Disconnection or Unreliable connection
- Low bandwidth

Screen Constraints

- Limited screen and small display
- Limited display and resolution
- Variable width and height ratio

H/W & S/W Constraints

- Limited processing capabilities
- Limited interaction techniques
- Limited input peripherals
- No full keyboard
- No mouse
- Limited memory and storage capabilities
- Device variability
- Limited graphical components and graphical libraries
- Complex and diverse data format
- Limited battery power or battery consumption

Mobility Constraints

- Auditory environment
- Visual environment
- Level of attention or concentration

They are watching

Data mining and surveillance has been an issue of debate for a long time. It is especially related to mobile devices because they are the main tool by which the surveillance is carried out. This issue comes to view when we deal with matters of national security.

Many times the ethics of surveillance is put to question by a lot of people. Today with the enormous usage of mobile devices it is easier for someone to get into the privacy of others. Sometimes the country or state does these things on the basis of security, especially in times of a terrorist attack, but we cannot know for sure that we are not being watched other times also.

The Snowden affair is the best example of why so many people are debating about surveillance. From all the fuzz we learned that the largest IT companies such as Microsoft, Apple and Google have been put to pressure by the government to give information about particular people. This directly has to do with mobile devices and technologies knowing that Microsoft has the Windows Mobile Platform, Apple has its iOS and of course Google with Android. In other words this has to do with the ethics.

Oscar H. Gandy makes a very nice description of how data mining and surveillance can be misleading based on the needs of the government. He has this to say:

“Data mining systems are designed to facilitate the identification and classification of individuals into distinct groups or segments. From the perspective of the commercial firm, and perhaps for the industry as a whole, we can understand the use of data mining as a discriminatory technology in the rational pursuit of profits... People tend to be outraged when they discover, or are informed that they have been discriminated against. There is some value, therefore, in supplying the press with egregious examples of individuals, or communities, or classes of people, who have been victimized by data mining, and by the use of profiles based on irrelevant attributes like

race or ethnicity. Of course, it is hard to predict how the public will respond to a growing awareness that discrimination is widespread, and that they are clearly at risk. In part it is a question about what people believe to be fair, and what they believe the circumstances demand. Unfortunately, I remained concerned that the use of data mining in the so - called “ war against terrorists” will soften us up for its use in the war against global competitors, or against the threat to shrinking profits, and a few “horror stories” about some “so-called victims” of discrimination will do little to shift the tide.”(7)

Mobile, data mining and the social web

The social web is an environment where people communicate the most. It is a vast area of data that could be mined. The social web mainly comprises Facebook, Twitter and LinkedIn and considering the large number of accounts on those sites and the huge multimedia exchange between users, data mining is a must, if managers of those sites want to have an overall control upon them. What is more important is the fact that these sites are mainly accessed by mobile devices.

“As humans, what are some things that we want that technology might help us to get?

- We want to be heard.
- We want to satisfy our curiosity.
- We want it easy.
- We want it now.

In the context of the current discussion, these are just a few observations that are generally true of humanity. We have a deeply rooted need to share our ideas and experiences, which gives us the ability to connect with other people, to be heard, and to feel a sense of worth and importance. We are curious about the world around us and how to organize and manipulate it, and we use communication to share our observations, ask questions, and engage with other people in meaningful dialogues about our quandaries.”(8)

This phenomenon that we call human nature, which means being curious, wanting to communicate and wanting to know more, which are the driving forces of the social web, is the main source for generating more data which need to be transformed into information and finally into knowledge, something that needs data mining.

Conclusion

In this paper we have tried to present the most important issues regarding data mining and mobile technologies. We learned that both of these fields are closely connected with each other and together they provide undisputed tools and methods in managing the large amounts of data that decision makers, managers and end users have to face on the IT realm. No matter if we use a cell phone, a tablet, a laptop or a smartwatch, we are part of a large worldwide network of data mobility and that is why it is very important to have a general overview of the relation between these two fields. We have tried to explain the importance of mining mobile devices and mining from mobile devices; this is a two-way relationship with one goal; achieving better results for greater success.

References

1. Data Mining on Mobile Devices for Personalisation!; Sam Tardif; University of Sydney; 2010
2. Framework for Interactive Data Mining Results Visualization on Mobile Devicesp; Muzammil Khan, Ali Shah, Israr Ahmad; Department of Computer Science & Software Technology University of Swat, KP Pakistan; 2014
3. Data Mining for Mobile Devices Using Web Services; R. R. Shelke, Dr. R. V. Dharaskar, Dr. V. M. Thakare; International Conference on Industrial Automation And Computing; 2014
4. Big Mobile Data Mining: Good or Evil?; Mirco Musolesi; University of Birmingham, Beyond Wires; 2013

5. Mobile Data Mining on Small Devices Through Web Services; Domenico Talia and Paolo Trunfio; University of Calabria; 2008
6. http://www.academia.edu/9883260/Overview_on_Data_Mining_Schemes_to_Design_Business_Intelligence_Framework_for_Mobile_Technology
7. Data mining and surveillance in the post-9.11 environment; Oscar H. Gandy, Jr.; Annenberg School for Communication, University of Pennsylvania; 2002
8. Mining the Social Web (Second Edition); Matthew A. Russell; O'Reilly Publishing; 2014