European Journal of Business and Management
ISSN 2222-1905 (Paper) ISSN 2222-2839 (Online)
Vol.6, No.26, 2014
www.iiste.org

The Data Traffic and Data Warehouses Store Managing and Controlling

Dr. Asmahan M. Altaher
Applied Science University, P.O.box 166-11391
A_altaher68@hotmail.com

Abstract
The new technology provides a number of problems, such as protection and security, better switch between towers and networks, the accelerating rate of technology improvements. The research aimed to fully addressed the data traffic managing and controlling; in the service companies. The paper answer the question, how networks communication firms manage and control data warehouses, scalable data warehouse, and processing data traffic, also the majority of the data traffic online, which requires massive bandwidth. The researcher recommend that corporate networks should have control standards and characteristics stand out in data warehouses, and the corporate network should fully care of the efficient traffic data management in order to reduce the data traffics problem and security issues.
Keywords: Data traffic, data warehouses, data management, privacy, security issues.

1. Introduction

The volume of data generated daily is growing at an exponential rate. The growth is due in part to the proliferation of sensors and increase of those sensors. To distill meaningful information from this growing mountain of data is need advanced data analysis techniques, such as cluster analysis. Central problem is therefore asserting where data processed in reliable, transparent and responsible manner, this has been recognized as traffic data security problem. A Control stander is needed for dealing with personal data, as the purposes of processing have an impact on the allowed transfer's data (P.Yu.J.Sendor,G.Serme,and A.S.de Oliveira, 2012).

A good understanding of traffic data passing the network is required for both planned and ad-hoc tasks T.J.Hacker,F.Romero, and C.D.Carothers,(2009. Capacity planning and traffic matrix processing, whereas traffic engineering, load–balancing, and intrusion detection are often require real-time behaviors. (B,Li, J. Springeer ,G. Bebis , and M. Hadi Gunes 2013).
Spatially that it expected to begin the 4G in 2012 or 2013. And, of course, fully expand will take a long time in 2017, because this network is not an easy task .The users need to move to 4G mobile networks, and the volume of data extremely large(,Y. Chang and M . Mitzenmacher, 2005).
The paper interfused a framework to answered the following questions

1. What is traffic data and how the networked firms manage and control data?
2. What are the collaborative efforts are needed in order to have efficient traffic data management and maintenance?

The remainder of this paper is organized as follows. In first Section was, introduced framework and discuss - phase of this framework from reviewed the previous studies.

In Second Section, were about methodology, finally, the researcher results and concludes.

2. Research framework

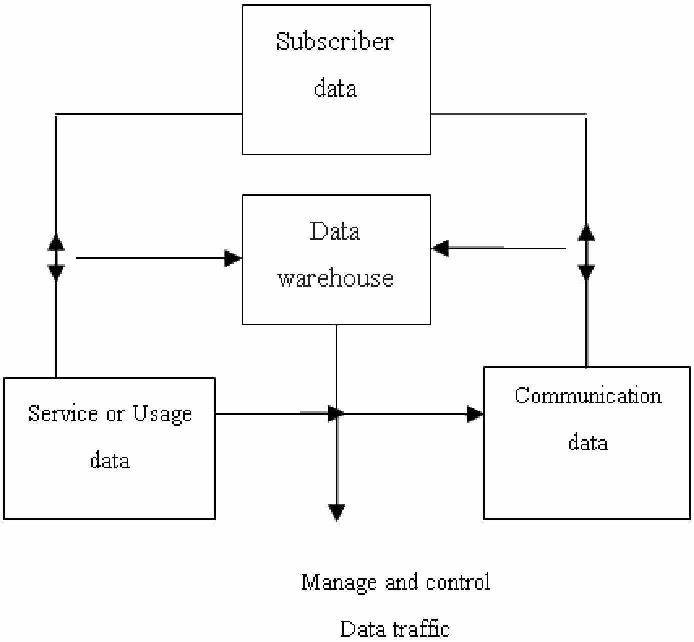

Figure (1) Research Framework Data Traffic

Resource: researcher adapted

Figure (1) represents a framework of the traffic data generally fall into four main categories.

Subscriber data provide a link between communication identifiers and a physical person. An LEA can use them as a phone book: given a network address or a phone number, the individual that owns it or the home it is connected to, along with other details that a provider might have can be retrieved.

Communication data provide a trace of who has talked to whom. Usually they link together communication identifiers, along with additional information such as the time and length of the connection or its status.

Location data provide information about where the two (or more) ends of a communication are physically located. This might include the GSM cell of a handset, or the base station of a wireless network connection.

Finally, Service or Usage data, contain information about the use that has been made of the network. While not logging the content of communications, information about the pages accessed on a web server, or the addresses to which an email was sent, can be considered as usage data (Fischer et al., 2010). The researcher adapts the four categories in order to study how to manage and control data traffic.

2.1 Previous studies

Traffic data is, from a technical point of view, an ill-defined term. Lobbying efforts in the U.K. have restricted the definition of traffic data included in the R.I.P. Act, only to include information identifying the physical piece of hardware used. Therefore only information about which physical server a user is accessing can be requested by LEAs. On the other hand, a recent document, leaked form EUROPOL, contains a wish list of traffic data to be seizable that encompasses all the log information that a web server would generate, down to the granularity of individual web pages(T. J. Hacker , F Romero , and C. D. Carothers,2009).

- Traffic usually means the amount of data that is either sent to or from the server on which your website is hosted.
- Traffic can also mean the number of unique visitors to, or individual (not necessarily unique) "hits", on your website (T. Benson . A, Anand, A. Akella. and M. Zhang, ,2009).Traffic will vary based on a number of factors, including:
- The popularity of your site and the number of visitors. For each visitor there is likely to be a small amount of ingoing data (when they send the request to retrieve your website content) and a larger amount (often around 10 times more) of outgoing data (when the content of your website is delivered to their computer via the internet).
- The services you provide with your website e.g. a web cam is perpetually uploading images/sound to your website for download by visitors. Services such as these will generate larger amounts of (data) traffic than a static site.

- The frequency with which your site is updated, remembering that every time you make changes to your website, you are uploading data to the website(DHS.,2010).

Furthermore traffic data are privacy sensitive because they can reveal private information that a user might not be aware of. The user is usually aware of the sensitive information contained and disclosed by content data, since it is explicitly communicated (D.G.Anderson and N. FEeamster.,2006). The same is not true for traffic data, since the information contained in them is extracted by a context that the initial user might not be aware of. For example, while it is obvious to most users who someone that intercepts a mobile phone call can overhear all conversation, it is not obvious that access to traffic data also reveals the location of the caller. It is even less obvious that the times, duration and frequency of calls can reveal data about the relationship of the communicating parties, although nothing about it has been explicitly revealed during the conversation (F. Bonomi, J. Zhu, and S. Addepalli, 2012). Traffic data of all sorts, already have a great amount of value because of the use one can make of them for direct marketing. Traffic data disclosure can also lead to individual embarrassment if they reveal practices that are not accepted as socially mainstream. The inability for the average person to model why their personal data is sensitive is often due to the fact that they do not perceive the threats to be more serious than leading to hassle or embarrassment (Google, 2013). This inability is often due to the fact that there is a lack of perception of oneself as an actor in society with plans and hopes that, for tactical reasons, might not be appropriate to be disclosed, in particular to persons that might not share them (George Danezis, 2010).

2.2 Data warehouse

The data warehouse stores event data in a plurality of related tables each associated with a particular event or event type. The tables include report-ready event tables that contain a subset, relating to a particular type of reporting, of the available event fields for a given event. Report generating applications that perform that particular type of reporting may then access all necessary fields for a given event in a single query, thereby increasing the performance of the system (Bouwman, E. Dommering, N. Van Eijk, N. Sitompoel, H. De Vlaam, 2004). Network system managers often require statistics and reports based on collected system events to alert them to system trends in order to proactively maintain normal system usage (Trajcevski , D., & Choudhary A., 2009). A data schema and method for storing network management system event data to allow efficient retrieval of information and aggregation for use in tracking system trends and generating useful reports is presented (Robinson, R.,(2006).

2.3 Data warehouse and traffic data

At the beginning of the twenty-first century, a rather specialized topic in the field of Telecommunications law began to receive a lot of attention: telecommunications traffic data. The context of this development was formed by the major technological changes which had occurred in the sector during the preceding decade. Two important innovations were introduced almost simultaneously in the 1990s: mobile telephony and the internet. Both communication technologies have undergone rapid developments and have now become part of everyday life. In their advance, new communication technologies have ousted wire line telephony. The traditional telephone network is being eroded at a fast pace; subscribers are trading in their fixed telephone subscriptions for mobile phones or switch to bundles where fixed telephony is part of a package of voice and data services (Google, 2013) .

These shifts in the telecommunications sector have caused new problems in various legal areas. These include issues surrounding the liberalization of the telecommunications market (competition and access), law enforcement on the internet (copyright enforcement, cybercrime and cyber-investigation) and privacy.

The subject matter of this study, communications traffic data, links with Data warehouse**,**
and all the issues related to. In short, communications traffic data are the resulted of data user communications (http://www.cisco.com) .

2.4 Efficient data management and Control

Minimum requirements are stipulated in GSM standards for the ability to operators to gather certain network management data. MSCs and location registers need to be able to generate data records for the following:
HLR interrogation;
- Location updating (HLR and VLR);
- Short Message Service (mobile originated, terminated, interworking MSC, GMSC);
- Common equipment usage.

The network needs to be able to generate these data in order to manage the database of subscribers and prepay users, and for technical network administration purposes (T.J.Hacker,F.Romero, and C.D.Carothers, 2009].Traffic data for managing GSM networks is quite technical, poorly standardized and virtually inaccessible for non-specialists. For this reason, the legal relevance of this category of traffic data remains limited

In order to fully efficient traffic management, and coordinated collaborative the following efforts are needed (A. J, Younge, R. Henschel, J. T. Brown, G. Von Laszewski, J. Qiu. And G, C. Fox. 2011):

(1) Develop new theoretical foundations that will: maximize the quality of the distributed information fusion from heterogeneous devices. Provide novel computational paradigms that will synchronize the decision and control process with the data generation process for the purpose of optimizing the traffic parameters (A. J, Younge, R. Henschel, J. T. Brown, G. Von Laszewski, J. Qiu. And G, C. Fox. ,2011)

(2) Develop novel methods, tools and components that will: enable seamless integration of the distributed data mining and control algorithms with the sensing and communication devices, at different level of spatio-temporal granularity and across heterogeneous networks.

(3) Provide coordinated query optimization and data management, for heterogeneous streaming data. Provide unified approaches to hardware/ software code sign for novel programmable sensing and communication devices Directive. (2002).
Figure (2) presents a global view of the main components that, we believe, should be.

Figure1. Data-Aware Traffic Control – System View pursued in a collaborative manner, in order to effectively address the problem of the efficient and effective traffic management. The top part of the figure, adapted from (Trajcevski et al., 2009) .illustrates the desiderata of the VII Consortium, towards the goal of efficient traffic management.

Positioning in this context, we seek to develop methodologies for context-aware management and mining of various spatio-temporal data, originating from heterogeneous sources and of different types, and use it to optimize the benefits of cooperative distributed control algorithms .

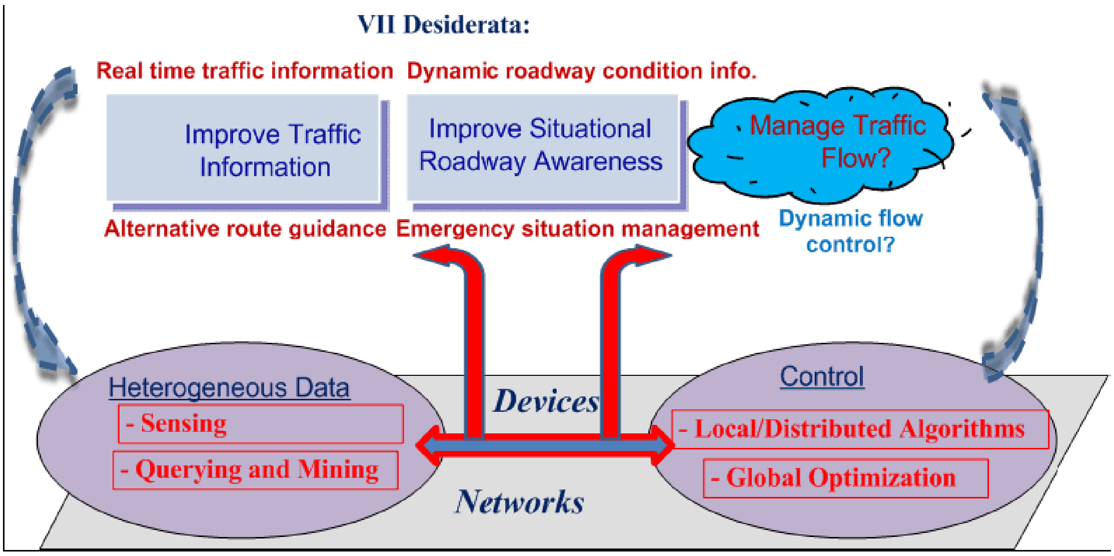

Figure (2). Data-Aware Traffic Control – System View
Resource (Trajcevski et al., 2009)
The next section of this paper is to dissect the Global traffic in mobile networks.

2.5 **Global traffic in mobile networks**

Figure 1 shows total global monthly data and voice traffic. It depicts a stable trend of data traffic growth with some seasonal variations. It shows that mobile data subscriptions will grow strongly, driving the growth in data traffic along with a continuous increase in the average data volumes per subscription. Highlights include:

- Data traffic doubled between Q1 2012 and Q1 2013
- Data traffic growth between Q4 2012 and Q1 2013 Was 19 percent

Voice traffic growth between Q1 2012 and Q1 2013 Was 4 percent

It should be noted that there are large differences in traffic levels between markets, regions and operators. These measurements have been performed by Ericsson over several years using a large base of commercial networks that together cover all regions of the world. They form a representative base for calculating world total traffic in mobile networks (N. Deng, Stewart, D. Gimach, M. Arlitt, and J. Kelley, 2013)

2.6 Examples of Mobile Data Traffic

Table (1) of Mobile Data Traffic Growth in (2013)

Region	**Mobile Traffic Growth Examples**
Korea	As reported by Korean regulator KCC, mobile data traffic on 2G, 3G, and 4G networks increased approximately 80% between January and November of 2012.
China	China Mobile's mobile data traffic grew 77% from mid-2011 to mid-2012. China Unicom's mobile data traffic grew 112% from mid-2011 to mid-2012.
Japan	As measured by Japanese regulator MIC, mobile data traffic grew 113% from September 2011 to September 2012.
Australia	As reported by Australian regulator ACMA, mobile data traffic grew 40% from mid-2011 to mid-2012.
Italy	As reported by Italian regulator AGCOM, mobile traffic in Italy in 3Q12 was up 32% year-over-year.
Global	Telefonica's total year-over-year mobile traffic growth was 35% in 1Q12, down from 75% in 1Q11. Vodafone's year-over-year mobile traffic growth was 34% in FY2012, down from 69% in FY2011.

Mobile data traffic will grow at a compound annual growth rate (CAGR) of 66 percent from 2012 to 2017, reaching 11.2 Exabyte's per month by 2017.

By the end of 2013, the number of mobile-connected devices will exceed the number of people on earth, and by 2017 there will be nearly 1.4 mobile devices per capita. There will be over 10 billion mobile-connected. Devices in 2017 including machine-to-machine (M2M) modules—exceeding the world's population at that time (7.6 billion).

Let us think of the following well-known image of communications: when user A Communicates with user B, a 'technical cloud' separates the two individuals: the network of the telecommunications provider. The services of the provider involve the transmission of the communication (the content, the actual message) from A to B. To achieve this goal, the provider uses certain technical data required for the proper functioning of the service. Examining the communications network from a legal point of view, there a need for a division of the standardized data warehouse: one part belongs to the user (the communication content) and one part belongs to the provider (the technical data necessary for the communication). The communication content belongs to the user, yet it is what the provider must convey and it should pass through the network without changes or provider inspection. The provider should have no interest in the content of the communications. The communication content is confidential; it is only shared between the communicating parties A and B. The involvement of the intermediate service provider is limited to the data transfer, the bare transport service. The communication content goes through the hands of the provider, but he remains an outsider, totally neutral and preferably invisible to A and B. In the network of the provider there also is, besides the communication content, data warehouse needed in the communication *process*. It concerns a broad category of data the provider must deal with concerning the technical and administrative functions of the network (A. Arcuri and L. Briand ,2011). Next the research design.

3. Research design

The consequence of this study is that entire the range of traffic data and how the networked firms manage and control and collaborative researches efforts are needed in order to have efficient traffic data management, also the relationships between the digital data cloud and traffic data . the researcher adapt the descriptive study, in order to placed exclusively within the context of horizontal traffic data studies, to share and disseminates common concepts related to the field of managing and controlling data traffic to contributes of the traffic data domain. Next result and conclusion.

3.1 Result and Conclusion

In order to solve the traffic data problem, the corporate network should fully care of the efficient traffic data management, the researcher recommends the below important points.
1- Data mining has the potential to provide significant insights both in terms of formulating strategies and policies as well as for actionable patterns, rules and predictive models that can be used to adapt and enhance real-time traffic management in a manner that will
2- The different system components, should trying to select a substitute technology to protect again serious problems and consideration of the technology (Trajcevski, G. et al., 2007).

3- Usually, public access phones, or terminals in libraries and cyber-cafes, offer an easy and relatively anonymous way of communicating. Traditional mail systems also provide less traffic data then its technological counterparts.

4- Another generic solution is to use proxies to access services. These can be used for web-browsing, mail, and can be implemented for mail by first sending it to the central offices of an organization before it is dispatched to its final recipient (N. Deng, Stewart, D. Gimach, M. Arlitt, and J. Kelley, 2013)
5- The firms should provide a standardized system format for user interfaces, along with a method for populating it. The data is easily accessible via standard interfaces to statistical process control, trend tracking, and report generating applications (Cohen, B. ,2003) .
6-Encryption can be used to hide some usage data, but not to generally protect traffic data.

References

F. Bonomi, J. Zhu, and S. Addepalli (2012) "Fog Comuting and its role in the internet of things " , in proceeding of MCC workshop on mabile computing conferences . pp13-26.

George Danezis,(2010) Comments on the E.U. Cybercrime,forum ttp://www.cl.cam.ac.uk/~gd216/RetentionComments.pdf

D.G.Anderson and N. FEeamster.(2006)" Challenges an opportunity in internet data mining" parallel Data Laboratory .Carnegie Mellon university.

T. J. Hacker, F Romero , and C. D. Carothers (2009)" An Analysis of Clustered failures on large supercomputing system " J, Parallel Distrib . comuting Vol . 69, no .7, pp. 652-665.

Trajcevski, D., & Choudhary A., (2009) Data-Aware Control Efficient Traffic Management (Extended Abstract)

Robinson, R.,(2006) VII Strategies for Safety and Mobility, in AASHTO RAC Annual Meeting 2006.

DHS. (2010) " Privacy Impact Assessments: Office of the privacy commissioner of Canada. Available http://www.priv.ge.ca/resource /fs-fi.

Directive. (2002)/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) , OJ L 201, Art. 5 (Confidentiality of the communications).

Cohen, B. (2003) Incentives Build Robustness in BitTorrent, http://www.ittc.ku.edu/~niehaus/classes/750-s06/documents/BT-description.pdf (May 2013).

Bouwman, E. Dommering, N. Van Eijk, N. Sitompoel, H. De Vlaam,(2004) Interconnectie: het vaste telefoonnet, het mobiele net en internet, Den Haag: SDU Uitgevers .

T. Benson . A, Anand, A. Akella. and M. Zhang, (2009)" Understanding Data Center Traffic characteristics " proceedings of SIGCOMM workshop: Research on enterprise networks. Pp 65-72.

Trajcevski, G. et al.(2007) Dynamics-Aware Similarity of Moving Objects Trajectories ACM-GIS. 2

A. J, Younge, R. Henschel, J. T. Brown, G. Von Laszewski, J. Qiu. And G, C. Fox. (2011) "Analysis of virtualization technology for High performance computing Environments ", IEEE international conference on Vol , 0. Pp 9-16.

Fischer et al., (2010) Communications Network Traffic Data Technical and Legal Aspects. A catalogue record is available from the Eindhoven University of Technology Library ISBN: 978-90-386-2339-9

Google, (2013) Cisco and/or its affiliates. All rights reserved. This document is Cisco Public. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525.
http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

N. Deng, Stewart, D. Gimach, M. Arlitt, and J. Kelley, (2013)"a adoptive green hosting " proceeding of international conference on Autonomic computing . California.ASU .

A. Arcuri and L. Briand (2011)" A practical guide for using statistical tests to assess randomized algorithms in software engineering " Proceeding international conference on software Engineering.

B,Li, J. Springeer ,G. Bebis , and M. Hadi Gunes (2013)" A survey of network flow application " Journal of network and computer application . Vol. 36. No ,2, pp 567-581.

T.J.Hacker,F.Romero, and C.D.Carothers,(2009)" An analysis of clustered failuier on large supercomputing system" J.Parallel istrib.Comput., Vol.69, no,pp ,652-665.

Google, (2013) Google compute Engine instances" https:// developers google .com/compute/ does/quota.

P.Yu.J.Sendor,G.Serme,and A.S.de Oliveira," (2012)Automating privacy enforcement in cloud platforms" in Dpm/STOP.Damiani,Vol .7731. Springer, pp160-173.

Y. Chang and M . Mitzenmacher (2005)" Privacy preserving keywords searches on remote encryption "in proc. ACM.

Author: Dr. Asmahan Altaher. Associate professor in Faculty of Business and Administration. Department /Management Information system .Applied Science. Private University. B.o.box 166-11391
Amman- Jordan –from 2008-till now .