Monitoring IT Power Consumption in a Research Center: Seven Facts

Monitoring IT Power Consumption in a Research Center: Seven Facts

Antonio Vetro', Luca Ardito, Maurizio Morisio
Dipartimento di Automatica e Informatica
Politecnico di Torino
Torino, Italy
e-mail: name.surname@polito.it

Giuseppe Procaccianti
Dipartimento di Automatica e Informatica
Politecnico di Torino
Torino, Italy
e-mail: name.surname@studenti.polito.it

***Abstract*—We analyze the power consumption of several IT devices placed in a research center affiliated to our University. The data collection lasted about one year and the analysis let us identify: i) the average instant power consumption of each type of device ii) trends of the instant power consumption curves iii) usage profiles and their power consumption iv) energy savings obtained from a different use of resources. Our main finding is that software and usage typology could affect power consumption more than hardware.**

Keywords—Green Computing, IT Energy Management, Electricity Meter, Data Analysis

I. Introduction and Related Work

Over the years, the use of Information Technology has exploded and IT has also contributed to environmental issues: the total electricity consumption by servers, computers, monitors, data communication equipment, etc. is increasing steadily [5]. According to [3], the ICT sector is responsible for a value between 2% and 10% of the worldwide energy consumption. Therefore, it is necessary to improve awareness in the IT industry with regard to environmental problems, and this aspect should be considered by the academic point of view [6]: in fact, "turning on research universities into living laboratories of the greener future" [8], will permit to quickly develop best practices and to make them available to industry and society in general. As an example we cite the work of Chiaraviglio et al. [1]: they applied a fully automatic measurement that is able to scale and track the number of devices powered on in real time. This technique has been applied at our same university, Politecnico di Torino. They created PoliSave, a software to turn on/off a PC by connecting directly to a website. PoliSave is being extended to all PC in the Campus, with the goal of saving about 250,000 € per year from the University energy bill. The study we present here is instead focused on the analysis of power consumption data, and it is designed to find out usage patterns of IT devices' energy consumption and to identify situations in which there is a waste of energy. Pinckard and Busch [2] also collected data on devices, focusing on the after-hours power state of networked devices in office buildings: they showed that most of devices are left powered on during night, concluding that this is the first cause of energy waste. Usage analysis is a crucial step to optimize the energy consumption: this task is even more necessary within data centers where the number of computers is large. In this field, Bein et al. [4] tried to improve the energy efficiency of data centers: they studied the cost of storing vast amounts of data on the servers in a data center and they proposed a cost measure together with an algorithm that minimizes such cost.

In our analysis the number of pc is lower, but we observe data from a real case. The paper is structured as follows: in the following section, the context of our work, including instrumentation and research questions, is described, section III presents results and, finally, section IV provides conclusions and future work.

II. Context of the Analysis

One of the strategic goals of Politecnico di Torino is the green footprint cutting and related costs reduction. Managers know that whenever a change is needed, the first step is to figure out the current scenario in a quantitative way, that means to measure [7]. Starting from the indicators, it is then possible to find solutions, improve results and solve problems. Therefore, several measures should be present on the dashboard of the *green power manager*, one of them is the electrical power consumption of devices: for this reason, our University decided to install in several departments sensors to monitor the power consumption of rooms, lighting and conditioning systems, data elaboration centers and single IT devices such as servers, printers, switches. We were involved in the measurement process of such data in a research center affiliate to Politecnico di Torino, the Istituto Superiore Mario Boella (ISMB), and we present in this paper data collected and some facts found.

A. Instrumentation

The measurement of power consumption was done through a power monitoring system provided by an industrial partner. The system is composed by sensors inserted between the monitored devices and the electrical plugs to which they are plugged in. For entire sections of lighting and conditioning systems, instead, the sensor is applied directly on the conductor through a pincer. Both type of sensors can compute active and reactive power, voltage, current intensity, cos φ, with a desired sampling time (we selected ten minutes). Data collected by sensors are sent to a bridge through ZigBee, then the bridge forwards the data via Ethernet/Internet to the central servers. Data are then accessible on a web portal and can be exported to be analyzed. We monitor the active power consumption on the following devices of the ISMB research center:

- Three distinct servers:
  - Server 1 (from 22nd April 2010 to 23rd November 2010)
  - Server 2 (from 22nd April 2010 to 23rd November 2010)
  - Server 3 (from 7th May 2010 to 23rd November 2010)
- A printer (from 5th March 2010 to 23rd November 2010)
- The conditioning system "CED1", that is cooling the room where Server 1 and 2 are located (from 23rd November 2009 to 23rd November 2010)
- The lighting system in Server 3 Room (from 24th November 2009 to 23rd November 2010)

We list in Table I the characteristics of the three servers. Server 1 and Server 2 are both used as web servers: they host web sites of research projects, where researchers share documents and files. Server 3 instead is used both as web server and to perform graphical operations. The printer is HP Laserjet P3005dn , with an operational power supply of 600W and standby consumption of 9W. Unfortunately we do not have information on the conditioning and lighting system. Finally, we define "instant power consumption" the average power consumption consumed in the sampling unit time (ten minutes).

B. *Research questions*

Eight different research questions drive data analysis. Firstly we have a group of questions (*Overview*) that is very general, and aims at discovering what is the actual average instant power consumption of the equipment in the ISMB research center, that we suppose being the typical equipment of similar centers. Overview questions are listed below:

1. What is the average instant power consumption of the servers in the last year?
2. What is the average instant power consumption of the printer in the last year?
3. What is the average instant power consumption of the other equipment (light, conditioning)?

After that, we focus our analysis on the power consumption of servers, because we can reduce their consumption only understanding how and how much they consume. The first question that is raised up is whether the power consumption of the three servers is the same or not:

4. Are there differences between the servers instant power consumptions?

Assuming, from the exploratory data analysis, that the power consumption in the studied context is not following any well-known distribution, we answer question 5 performing the Wilcoxon Two Sample test [9]. The difference we try to find with this question is an inter-server difference: the next step is to explore the aspects related to the progress of the single servers' power consumptions. Initially, we investigate whether the power consumption is homogeneous or variant:

5. Are there any peaks in instant power consumption or is it homogeneous? If so, how long do they last? Are they relevant, in terms of power values?

We answer to this question in a qualitative way, i.e. plotting for each server 2 different graphs: the power consumption over time and the estimated probability density. The first plot let us identify the presence of peaks and trends in the observed time window, whilst the latter permits to see if power consumption accumulated by peaks is relevant, looking to the frequency of the values associated to the peaks. Peaks represent a rapid growth or decrease, or deviations from a normal behavior. However, a server could have several behaviors in terms of power consumption, associated for example to a different load or a particular software or hardware configuration. Therefore the scope of the next question is to understand the existence of different power consumption "behaviors", that we call "profiles".

6. Can we identify different usage profiles (e.g. active/standby)?

We perform a cluster analysis to answer question 7. We use the K-Means algorithm [10] and the bivariate plots, obtained through normalization and rescaling of the variables (watt-time couples). The selected clustering algorithm aims to group and find aggregations of data around certain values called "centers", which could point out different power consumption "profiles" and relationships between the variables (for instance, a typical profile in common computers is the standby profile).

TABLE I. SERVERS' TECHNICAL CHARACTERISTICS

	Server 1	**Server 2**	**Server 3**
Type	Dell PE r300	Dell PE1950 III	Dell Precision T5400
RAM	8 GB DDR 2	4 GB	4 GB
Proc	Quad Core Xeon X5460 3.16 GHz 64 bits	Quad Core Xeon E5410 2.33GHZ 32 bits	Dual Core Xeon 5200 2.49 GHz 64 bits
Power supply	400 W	670 W	875 W
Operating system	Windows Server 2003 R2 Enterprise X64	Ubuntu 2.6.24-19-server	i)Windows Server 2008 ii)Ubuntu 10.04 Server iii)Windows XP
Energy certification	NO	NO	Energy Star 4.0

The plots let us visualize and verify the clusters found. If profiles are found, it is also important to verify if they correspond to daily/nightly activities, relating, for each server, the progress of power consumption with time tables of human activities in ISMB. Hence, we plot, for each server, the power consumption in a whole day, selecting for each week of the last 3 months a random day between Tuesday, Wednesday and Thursday (we avoid week ends, Mondays and Fridays because typically in these days human activities are not representative of the typical work-day). Observing the 12 plots obtained and interviewing people working in the center, we are able to identify the time range in which the majority of activities on all the 3 servers is carried out, that is between 9 a.m. and 8 p.m. At this point, it is possible to divide data in daily and nightly consumption, and then compare the two subsets, using Wilcoxon Two Sample test, since data is not normally distributed.

Moreover, we are interested in pointing out the power consumption of each profile, in order to understand how much energy is saved/lost by applying the configurations and conditions that determine the different power profiles. This is done by tagging each observation with the profile it belongs to and then summing up the cumulated consumption. Therefore, the research question is:

7. How much total energy did servers consume in the last year in the different profiles?

Finally, the same question is replicated to the printer, that has two well-known profiles: an active profile when it is printing, and a stand-by profile when it is not.

8. How much energy can we save by turning off the printer when it doesn't work?

All questions are about power, and related data are expressed in Watt (W), exception given for RQ 7-8 that measure energy (KWh).

C. *Threats to validity*

The first threat of this research is an external threat: the analysis is performed on specific machines, thus generalizing these results is not possible. However, it can be possible to look at this equipment as representative of a category of equipments with similar characteristics.

Further, a derived internal threat is that the information on the technical characteristics of the IT equipment (printer, servers) and on their usage (massive, constant, etc) could be not enough to deeply motivate all the curves of the power consumption analyzed and determine with precision the impact on the measures. Therefore, the causes that we derive from the observation, can be biased.

Finally, we also identify a conclusion threat determined by the sampling time (ten minutes): as a consequence we have average values even for instant power consumption measures, and we could miss some fluctuations.

III. Analysis Results

A. *Overview(RQ1 to RQ3)*

We provide on Table II two descriptive statistics about RQ1 to RQ4: the average instant power consumption and the index of dispersion that quantifies how much data is sparse around the mean.

B. *Servers (RQ4 to RQ7)*

RQ4: Are there differences between the servers instant power consumptions?

We observe in Table III that power consumptions of Server 1 is very different from the consumptions of Server 2 and 3. However, Server 2 and 3, even if similar in mean values (difference is only 8 Watts), have statistically different mean power consumptions.

RQ5: Are there any peaks in instant power consumption or is it homogeneous? If so, how long do they last? Are they relevant, in terms of power values?

Looking to the time plots, we observe for Server 1 (Figure 1) many high spikes (that reach values that are more than the 50% of the mean), two low spikes and frequent switches between low and high values. However, the index of dispersion (see Table II) is reduced, that means that the time duration of the peaks is short. For Server 2 (Figure 2), data is more concentrated around the mean value, and the peaks (3 low spikes and a dozen of high peaks) lasts for short periods of time (but longer than Server 1, as the index of dispersion suggests).

TABLE II. Average instant power consumption of selected devices

RQ	Device	Average Instant Power Consumption	Index of dispersion (var/mean)
1	Server 1	108.02 W	0.55
1	Server 2	145.12 W	2.3
1	Server 3	139.63 W	22.44
2	Printer	13.46 W	199.95
3	Light (Server 3)	107.76 W	877.56
4	Conditioning (Server 1+2)	2713.77 W	880.12

TABLE III. Result of Wilcoxon test on servers instant power consumptions

Comparison	95% Difference Confidence Interval	P- val	Different?
Server 1 vs Server 2	{ -37.78 , -37.68 }	< 2.2e-16	YES
Server 1 vs Server 3	{ -47.57 , -47.36 }	< 2.2e-16	YES
Server 2 vs Server3	{ -8.43 , -8.33 }	< 2.2e-16	YES

Figure 1. Instant power consumption over time (Server 1)

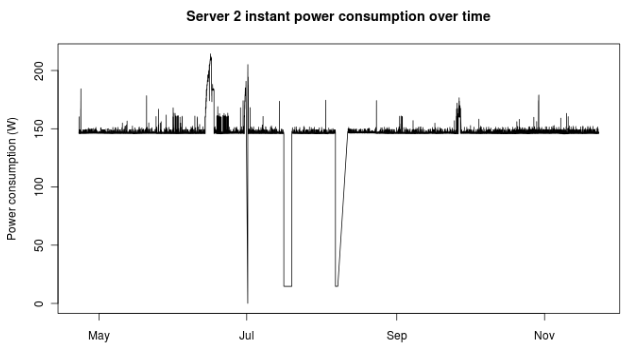

Figure 2. Instant power consumption over time (Server 1)

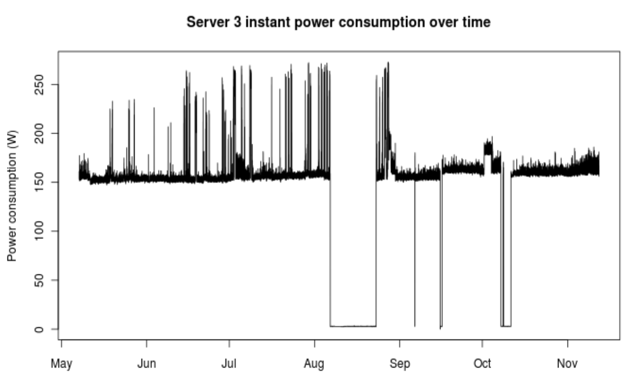

Figure 3. Instant power consumption over time (Server 3)

Finally, for Server 3 (Figure 3) the situation is yet different: it has a higher consumption and many high peaks until the end of August, then lower power consumption and peaks starting from September 2010. The change in the curve has a motivation: the server was used to perform continuous intensive tasks as image processing, parallel coding and massive video/audio streaming until end of August. Then, it was used as a normal web server, as Server 1 and Server 2.

The variability of data has also the same behavior: higher until September, then reduced. Moreover, there is a very long period (about 20 days) of zero power consumption (it was powered down), followed by 4 other smaller periods of zero consumption.

Plotting instead the probability density estimation of the servers' power consumptions, we can see around which values data is concentrated, and so if peaks are relevant both in terms of power consumption and duration. Server 1 (Figure 4) has 3 main concentrations of data: the highest is around the mean value (~108 W), then there is a similar peak at about 112 W and a lower one in their middle, finally two very low peaks at the two extremes of the graph. We conclude that peaks of Server 1 are relevant in terms of duration, but not in terms of variation from the mean value. The probability function of Server 2 (Figure 5) is totally different: data is concentrated around the mean value, and the distribution is very similar to a normal distribution with very low variance. The higher index of dispersion is due to the small peak toward the 10W and the other one on the right of the mean. We observe that peaks for Server 2 are irrelevant in terms of duration, but some are quite far from the mean. Finally, we observe Server 3 in Figure 6: except the peak around the mean, the 2 big peaks (~0 W and ~160 W) have high probabilities, whilst the small peak on the right is quite far from the mean. As a consequence, spikes of Server 3 are relevant both for time length and power consumption. This concludes the answer of RQ6.

RQ7: Can we identify different usage profiles (e.g. active/standby)?

We obtain from the K-Means algorithm 5 clusters for Server 1, 4 clusters for Server 2 and 4 Clusters for Server 3. The centers of the clusters are the following, in increasing order of power (W) :

- Server 1: 8.36, 105.23, 106.97, 109.44, 110.47
- Server 2: 14.64, 146.44, 146.68, 149.22
- Server 3: 2.80, 154.51, 160.56, 246.81

We can surely identify a “low power profile” for all the servers (the lowest-value center). Instead active profiles could be more than one, especially for Server 1 (data varies from 105 to 110 W) and Server 3 (where the difference is clear, since values go from 154 to 246 W). For this reason, we perform a further cluster analysis (Figure 7), focused just on the active profile, that allow us to gain more information. We find the following centers:

- Server 1: 104.93, 105.21, 105.92, 109.76, 111.10, 138.61
- Server 2: 146.46 , 146.63, 185.10
- Server 3 : 154.27, 160.64, 245.87

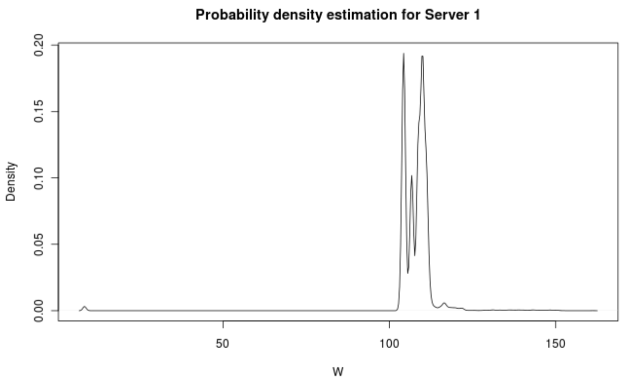

Figure 4. Probability density function estimation (Server 1)

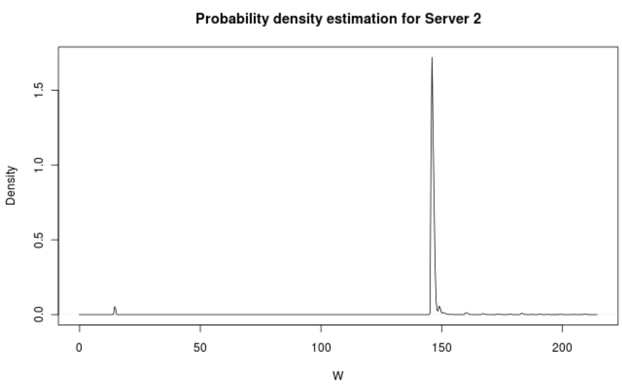

Figure 5. Probability density function estimation (Server 2)

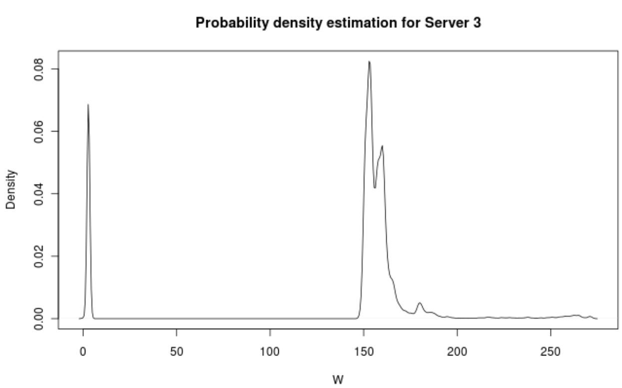

Figure 6. Probability density function estimation (Server 3)

Reducing to significant values, we can identify for all the servers at least two different high power profiles and a low power profile: we show values in Table IV . Subsequently, we investigate whether the two high power profiles are related to the day/night human activities. Even if the Wilcoxon statistical tests (Table V) verifies the difference between daily and nightly power consumption with the standard level of confidence of 95%, this difference is in practice negligible, since it's in the order of μW for Server 1 and Server 2 and it is less then 1 W for Server 3.

RQ8: How much total energy did servers consume in the last year in the different profiles?

The estimated cumulated energy consumption of servers in the different power profiles is shown in column "kWh" of Table VI, whilst the column "%" shows the percentage of each cumulated power with respect to the total.

C. *Printer (RQ8)*

RQ9: How much energy can we save by turning off the printer when it doesn't work?

This is the overall data of the energy consumption of the printer in the Active – Standby profiles:

- Active: 11.93 kWh
- Standby: 21.29 kWh

The difference is important indeed, 2/3 of energy are used (wasted) in the standby mode.

TABLE IV. SERVERS' POWER CONSUMPTION PROFILES OBTAINED THROUGH CLUSTERING OF DATA

Servers	**High power profile 1 (W)**	**High Power Profile 2 (W)**	**Low Power Profile (W)**
Server 1	~138	~105	~8
Server 2	~185	~146	~14
Server 3	~245	~160	~3

TABLE V. SERVERS' POWER CONSUMPTION COMPARISON BETWEEN DAY AND NIGHT

Server	**95% Diff C.I. Between Day/Night (W)**	**P- val**	**Different?**
Server 1	{-1.00 e-05, -1.80 e-05}	0.05	YES
Server 2	{3.86 e-05 , 4.22 e-05}	0.01	YES
Server3	{ 0.625 , 0.937 }	< 2.2e-16	YES

TABLE VI. CUMULATIVE POWER CONSUMPTION BY PROFILES

Servers	**High power profile 1**		**High Power profile 2**		**Low power profile**	
	kWh	**%**	**kWh**	**%**	**kWh**	**%**
Server 1	20.37	3.75	523.05	96.21	0.22	0.04
Server 2	33.57	4.62	692.2	95.17	1.57	0.22
Server 3	89.09	13.82	554.25	85.98	1.29	0.2

IV. CONCLUSIONS AND FUTURE WORK

We analyzed the power consumption in the last months of the equipment in the research center ISMB, affiliated to our University. We monitored at high level data about conditioning and lighting systems and general devices and we conducted a more detailed analysis on the servers and the printer. We draw from the statistical analysis and the answers of the research questions the following facts, and related further questions for future work.

Servers

Fact 1. We found differences between the power consumptions of the three servers (RQ4), likely determined by software usage, not by hardware equipment. In fact, despite Server 1 has a more powerful hardware equipment (CPU, memory), it has the lowest power consumption.

Fact 2. The power consumption of servers is not homogeneous over time (RQ5).

There are several peaks. Peaks are determined by software usage: as a matter of fact, Server 3 consumes up to 75% more when it is used for graphical operations.

Fact 3. Servers have different power consumption profiles (RQ7). This is determined by software usage.

Fact 4. Conditioning and lighting for servers consume more than computation (especially conditioning, that consumes approximately ten times more) (RQ3).

Fact 5. Low power profile (or Stand-by) for servers is useless (< 1%) (RQ7).

Fact 6. There is no substantial difference between day and night servers' power consumption (RQ7).

Printer

Fact 7. The printer consumed more energy in standby mode than in active mode (RQ9).

Indeed, in our analysis, with a mechanism able to turn off the printer when it doesn't work, 21 kWh would have been saved in the period March-November 2010, which is the 64% of the printer's total power consumption in that time range. Or, alternatively, shutting down the printer during night, it is possible to save the standby power consumption (13 W, despite the 9W declared in the technical sheet), that means to save 47.45 KWh per year.

Even if this analysis is very initial, and specific to few machines that may not be representative of the whole population, we believe it points out some simple checks that every energy manager should do as a first step to reduce energy consumption: consumption of conditioners and lighting, consumptions of printers in idle mode, consumption of servers both over day and night. Moreover, data we have presented could be compared to other analysis on equipment with similar characteristics.

Future work will be devoted to understand more the reasons of the behaviors observed. Notably to investigate deeper the motivations of the differences of servers power consumptions, and to verify them experimentally by setting up different configurations/conditions in the machines to evaluate their impact on power consumption. Secondly, we will repeat the same analysis on the Data Elaboration Center of our university, and we will compare it with the data presented in this paper. We think that the research questions that drove this analysis could be adopted and answered by other researchers in different universities and centers: building up a common benchmark of power consumption, it is possible to identify common and efficient solutions that can be then exported in industry and society.

ACKNOWLEDGMENTS

Authors thank Antonino Fiume for his precious help in the handling with the monitoring system youmeter [11], and Maurizio Molinaro for his help in collecting information at ISMB.

REFERENCES

[1] L. Chiaraviglio and M. Mellia, "PoliSave: Efficient Power Management of Campus PCs", Telecommunications and Computer Networks (SoftCOM), 2010 International Conference on, Publication Year: 2010, Page(s): 82 - 87

[2] J. A. Roberson, C. Webber, M. McWhinney, R. Brown, M. Pinckard, and J. Busch, "After-hours Power Status of Office Equipment and Energy use of Miscellaneous Plug-load Equipment", Report LBNL-53729-Revised, LBNL, Berkeley, CA, Pubblication Year: 2004

[3] SMART 2020 Report, "Enabling the low carbon economy in the information age", http://www.theclimategroup.org

[4] D. Bein, W. Bein, and S. Phoha, "Efficient Data Centers, Cloud Computing in the Future of Distributed Computing", Information Technology: New Generations (ITNG), 2010 Seventh International Conference on, Publication Year: 2010, Page(s): 70 - 75

[5] S.Herat, "Sustainable Management of Electronic Waste (e-Waste)", CLEAN – Soil, Air, Water, Volume 35, Issue 4, Publication Year: 2007, Page(s) 305–310

[6] Y. Wati and K. Chulmo, "The Green IT Practices of Nokia, Samsung, Sony, and Sony Ericsson: Content Analysis Approach", 43rd Hawaii International Conference on System Sciences (HICSS), Publication Year: 2010, Page(s): 1 - 10

[7] T. DeMarco, "Controlling software projects: management, measurement & estimation", ISBN: 0917072324, Publisher: Prentice Hall, Publication Year: 1986

[8] D. Estrin, "Participatory sensing: applications and architecture [Internet Predictions]", IEEE Internet Computing, Volume 14, Issue 1, Pubblication Year: 2010, Page(s) 12-42

[9] L. Sachs, "Applied Statistics - A Handbook of techniques", ISBN: 0387909761, Publisher: Springer, Publication Year: 1978

[10] P.Tan, M. Steinbach, and K. Vipin, "Introduction to Data Mining", ISBN: 0321321367, Publisher: Addison-Wesley, Publication Year: 2005

[11] https://www.youmeter.it/youmeter/ , last access 7th March 2011