International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2014): 5.611

Context Awareness in Data Mining Applications – A Survey

Sanjeevani S. Dhaneshwar[1], Manisha R. Patil[2]

[1]Post Graduate Student at Savitribai Phule Pune University, Smt. Kashibai Navale College of Engineering, Vadagaon Budruk, Maharashtra, India

[2]Assistant Professor, pursuing Ph.D. at Savitribai Phule Pune University, Smt. Kashibai Navale College of Engineering, Vadagaon Budruk, Maharashtra, India

Abstract: ***With the advent of the digital explosion and big data, data mining and analysis play a key role in any business development. Tremendous amount of data is collected by organizations through their day to day activities and transactions – both internal and external. There is an intrinsic relationship between data items as a result of the effects of the surrounding situation or environment in which they are produced. This relationship can be leveraged to search out attention-grabbing patterns and trends within the data. Such systems are called "context aware" systems. This paper surveys the various approaches and applications that add the aspect of context awareness to data mining applications.***

Keywords: context awareness, data mining.

1. Introduction

As the world is getting more and more connected, the number of online interactions an average user has with other users or applications has increased exponentially. A huge amount of data about these interactions is generated as a result. It acts as raw material to extract business-essential knowledge by analyzing the data. The business advantage of adopting this approach is obvious.

Over the years a number of techniques and algorithms have been developed and implemented for effective knowledge discovery. The results of data mining can be improved further by exploring the relationship that is implicit within the data. Data generated as a result of any action is always influenced by the surrounding factors. The object producing data is said to be operating in a "context". Adding context awareness to data mining processes has been extensively used to improve their accuracy.

In the paragraphs to follow, firstly, the concept of context is described in detail. It is then followed by some approaches to implementation of context aware data mining applications.

2. The Notion of Context

Adomavicius et al. (2011) describe the notion of context as a frame for a given object. This frame contains a set of elements that influence the object and its actions [1]. For example, a lady buying books may search for bedtime stories for her kids, but may search for recipe collections when buying for herself. The factor influencing her actions in this case is her relationship with the person she wants to buy the books for. Other examples of contextual factors are location, time of the day, day of the week, month, purpose of purchase, etc. Figure 1 provides an illustration of this concept.

Some contextual factors are fully observable, i.e. the possible values that they may take and their structure is already known. For example a month is made up of days, which are made up of hours, which in turn are made up of minutes and so on. Sometimes all the factors to take the optimum decision are not known. These are called partially observable. In some cases, no explicit information about the context parameters is known. It has to use latent knowledge of context implicitly. They are called un-observable contextual factors.

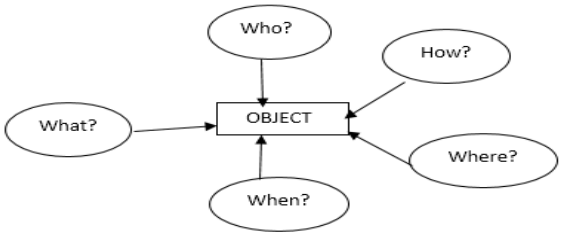

Figure 1: Context – The factors influencing an object

A system must also take care of distinguishing between those factors that will change over time and those that won't. Factors like country (location) probably won't change over a long period of time. But factors like purpose of purchase, time, etc. will frequently change.

3. Approaches to implementing context aware data mining systems

3.1 Using Context in different stages of algorithm

Context awareness has a very useful application in personalized recommendation systems. Recommendation systems analyze user behavior, history and preferences to compile a list of items (for shopping, browsing, etc.) that the user is most likely to be interested in. Adomavicius et al. (2011) describe three approaches, based on how the contextual information is used: [1]

- *Contextual prefiltering*: Contextual information is used as an input for selecting the items for recommending. Then the actual recommendation is performed using any 2D recommender algorithm.
- *Contextual postfiltering*: Contextual information is used at a later stage. First, user preferences and the ratings are predicted using any traditional 2D recommender system on the entire data. Then, the resulting set of recommendations is modified for each user using the contextual information.
- *Contextual modeling*: Contextual information is used as a part of the modeling process for the estimation of user ratings. These user ratings are then used further in the recommendation process.

3.2 Context as a means of Self Supervised Learning

Yuan and Wu (2008) have an interesting take on the learning methodology adopted in context aware clustering. Clustering is an unsupervised form of learning. It generally uses distance metrics that measure similarity or dissimilarity of data items. Data is grouped together such that items within a group are similar to one another.

According to Yuan and Wu, as it is generally quite difficult to choose the right distance metric in advance, we tend to learn a good metric by imposing supervision [2]. They recommend that supervision can come from the inside rather than as an external input, wherein the contextual relationships between the data serve as the supervising factor. This approach is called as self-supervised clustering.

3.3 Context Weighting

Context may vary from user to user. Chen et al. (2015) have developed a framework, ConRec that uses a context weighting method to find out which factors are most influential to user's decision [3]. ConRec computes context weight for each user by calculating the fluctuation coefficient. This coefficient represents the changes in a user's rating within a type of context. The computed context weights are then represented as a User – Context Matrix as shown.

Table 1: User Context Matrix

	C_1	…	C_m
$User_1$	$Weight_{11}$	…	$Weight_{1m}$
:	:	…	:
$User_u$	$Weight_{u1}$	…	$Weight_{um}$

Consider an *n*-dimensional space, where *User* and *Item* are 2 dimensions. The remaining *n-2* dimensions constitute the context. This multidimensional data space is decomposed to multiple 3D – spaces. Each space contains 3 dimensions – user, item and one type of context. The rating function then used for recommendation can be defined as:

$$R: U \times I \times C_1 \times C_2 \ldots C_n \rightarrow rating \quad (1)$$

3.4 Contextual Grouping

Wallace and Stamou (2002) investigate the mining of user interests in usage of media documents using a context aware hierarchical clustering algorithm. They suggest that a user may be interested in a particular media document group for a number of *reasons*. For example, a user may like a particular collection of movies because of the actors playing in them. Similarity of dissimilarity measures relating to some reason can be defined on each group. E.g. similarity between actors. These are called *relations*.

If *m* different relations r_1, r_2 ..., r_m, are defined, then, for every pair of clusters, *m* different compatibility indices (CIs), f(C1,C2, ,r_1,),f(Cl,C2 ,r_2)....,f(C1,C2,r_m) can be defined, where a compatibility index indicates the quality of the cluster. Each relation can be considered as a context. Each CI can be considered as an indication of cluster similarity between the two candidate clusters, as far as the corresponding context is concerned [4].

3.5 Context as social relationships

Stan et al. (2014) describe the importance of incorporating social networks into recommendation systems. They define social networks to be modelled as graphs, where entities are described as nodes and their relationships through links. Social networks are characterized by: [5]

- Users (with a user profile, activities and connections),
- Social objects (topics of user interactions, shared videos, photos, etc.),
- Social connections (people belonging to different communities or social spheres)

The aspect of social connections includes two concepts of *trust* and *reputation*. A person is more likely to go for recommendations that are liked by his or her friend because the person shares a trusted relationship with that friend. Also it is seen that reputed personalities (public figures) also tend to influence the choices of the mass. Other sources of information from social networks are user profile, user interests obtained from the posts liked or shared by the users or the statuses and comments posted by them.

3.6 User Centered Context

Traditionally, sensors are used to collect information about the context and its interpretation. However Mowafi and Zhang (2007) have made use of the natural, intuitive interaction that users have with their devices to extract context. This approach is called the user centered approach. It is more interactive and more actively driven by user's interests. Preference is given to user dialogue than to the situational context. Thus, users gradually build their own context. Programming-by-demonstration [7] approach is used for this. Context learning is performed by allowing users to annotate models of context through examples wherein they express their preference in that particular context.

The following two studies present 2 domains that have embedded context into their processes.

3.7 Context in the Learning Domain

Globalization of businesses has led to e-learning becoming a crucial part of any organization. It uses an online Learning Management System that allows that assigns learning courses or modules to users, allows user to take courses, and tracks the user's progress. Introducing context into this domain has improved the quality of learning as it can be made more personalized. Zheng et al. (2009) have explored three types of context awareness in the e-learning domain: [6]

- *Awareness to knowledge context:* This refers to the user's skill set, experience, training requirements and factors that will have direct effect on the effectiveness of learning provided.
- *Awareness to social context:* This includes the interpersonal relationships between the users. It is useful for knowledge transfer and collaborative training efforts.
- *Awareness to technical context:* Each type of media used to provide learning (like audio, video, text or graphics) has a different impact on the learner's mind. This fact can be exploited to enhance training quality by studying the characteristics of the user (age, experience, domain of work) and using the right combination of media to maximize learning.

3.8 Context in Collaborative Applications

The last study to be presented in this paper, provides the big picture of how context is used in collaborative applications. Fadzillah et al. (2012) identify seven context awareness entities; application, media, method, tool, platform, framework and device. These entities work with eight context aware elements - domain, activity of user, context object, locations, type of communication, type of context, digital elements and models [8]. The eight elements can be clustered into three main categories:

- user (people),
- resource (peripherals)
- task (activities).

Their findings conclude that contextual information in a collaborative environment, such as who is online, what they are working on, how much they have progressed in their given task, and so on, is vital to make timely and accurate decisions.

3.9 The 2 views of context

Context can be viewed in ways [1] –

- Representational view: The attributes of context are known prior to its application to the computation problem. Also, these attributes tend not to change.
- Interactional view: The context features change dynamically. The user activity leads to the formation of a context, which in turn influences the user activity.

Some of the studies presented above may be roughly categorized into either of the views as follows:

Table 2: Context Views

Type of View	Study Reference Number
Representational	[2],[3],[4],[6]
Interactional	[5],[6],[8]

As we can see in Table 2, [6] falls in both views, as it defines three different types of context. The social context defined by it may be interactional in nature, whereas the remaining two types may be representational.

4. Conclusion

This paper has presented a list of varied applications and approaches through which context awareness is being practically implemented. Context, combined with data mining and analysis, plays a key role in enabling ubiquitous computing. It is an upcoming trend that has taken a step further into unifying all domains of technology. We hope that the findings summarized in this paper will help the readers to grasp the general concept of context awareness and pique their interest into exploring it.

References

[1] Gediminas Adomavicius, Bamshad Mobasher, Francesco Ricci, Alex Tuzhilin, "Context Aware Recommender Systems", AI MAGAZINE Fall 2011. ISSN 0738-4602.
[2] Junsong Yuan, Ying Wu, "Context-Aware Clustering," In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. ISSN 1063-6919
[3] Bin Chen, Ping Yu, Chun Cao, Feng Xu, Jian Lu, "ConRec: a Software Framework for Context-aware Recommendation based on Dynamic and Personalized Context," In 2015 IEEE 39th Annual International Computers, Software & Applications Conference
[4] M. Wallace. G. Stamou, "Towards a context aware mining of user interests for consumption of multimedia documents", In Multimedia and Expo, 2002. ICME '02. Proceedings. 2002 IEEE International Conference on
[5] Johann Stan, Fabrice Muhlenbach, Christine Largeron, "Recommender Systems using Social Network Analysis: Challenges and Future Trends", Alhajj,Reda, and Rokne, Jon. Encyclopedia of Social Network Analysis and Mining, Springer, pp.1-22, 2014, 978-1-4614-6171-5
[6] Yanlin Zheng, Luyi Li, Fanglin Zheng," Context-awareness Support for Content Recommendation in e-learning Environments," In 2009 International Conference on Information Management, Innovation Management and Industrial Engineering
[7] Mowafi, Y, Dongsong Zhang, "A User-centered Approach to Context-awareness in Mobile Computing," In Mobile and Ubiquitous Systems: Networking & Services, 2007. MobiQuitous 2007. Fourth Annual International Conference on
[8] N. A. A. Fadzillah, N. Omar, S. Z. Z. Abidin, "Application-based Context-Awareness in Collaborative Workspaces: A Review", In 2012 IEEE International Power Engineering and Optimization Conference (PEOCO2012), Melaka, Malaysia: 6-7 June 2012