© April 2018 | IJIRT | Volume 4 Issue 11 | ISSN: 2349-6002

A Situational Analytic Method for User Behavior Pattern in Multimedia Social Networks

G Rajendra[1], G.V. Ramesh Babu[2]
[1]*Student, Dept. Of Computer Science, SVU College of CM & CS, Tirupati*
[2]*Assistant Professor, Dept. Of Computer Science, SVU College of CM & CS, Tirupati*

Abstract- **The past decade has witnessed the emergence and progress of multimedia social networks (MSNs), which have explosively and tremendously increased to penetrate every corner of our lives, leisure and work. Moreover, mobile Internet and mobile terminals enable users to access to MSNs at anytime, anywhere, on behalf of any identity, including role and group. Therefore, the interaction behaviors between users and MSNs are becoming more comprehensive and complicated. This paper primarily extended and enriched the situation analytics framework for the specific social domain, named as SocialSitu, and further proposed a novel algorithm for users' intention serialization analysis based on classic Generalized Sequential Pattern (GSP). We leveraged the huge volume of user behaviors records to explore the frequent sequence mode that is necessary to predict user intention. Our experiment selected two general kinds of intentions: playing and sharing of multimedia, which are the most common in MSNs, based on the intention serialization algorithm under different minimum support threshold (Min_Support). By using the users' microscopic behaviors analysis on intentions, we found that the optimal behavior patterns of each user under the Min_Support, and a user's behavior patterns are different due to his/her identity variations in a large volume of sessions data. Finally, we identify some common properties of the most popular multimedia content.**

1. INTRODUCTION

The rapid development of Multimedia Social Networks (MSNs) causes the tremendous growth of users and digital contents. It's also convenient for users to access digital contents in MSNs with a large-scale video dataset. Meanwhile, the interaction between user and user, user and system increases. Therefore, providing users with timely and rapidly personalized services considering the complex interaction is now a challenge in the study of multimedia social networks. Generally speaking, multimedia computing can be decomposed into three different stages, from data-centric multimedia compression, content-centric multimedia communication and content analysis, to user-centric social media analysis till today, including user trust modeling , propagation paths mining and digital right sharing , and digital forensics. However, understanding and predicting what multimedia content users' real needs in different situations and contexts have not been well studied. Context-Aware (CA) was first proposed by Schilit in 1994. They defined context as the set of location, people nearby, objects, and the changes of the objects. Prof. Carl K. Chang proposed the Situ theory by combining the service environment with situation awareness to handle the dynamic update or development of service at run time. Therefore the service can meet the changing needs of users and provide users with personalized service. In order to adapt to the dynamic service environment and make a timely respond to the feedback of service environment, social media services increasingly require situation awareness. In social media networks, the human being is a complex and open system. The individual's intention can change at any time, which also causes a change in the user's needs. Moreover, the user's context and characteristics of the dynamic change will have different effects in a user's potential needs . A user's intention can be reflected through the acquiring attributes of the user's situation awareness and feedback on resources. The system can formulate a timely personalized service for the user based on user's intention, which will increase the user's service experience.

In social media networks, the user has different roles in different groups. The different identifications that the user has may cause the user's intention to change. The change of intention reflects the change in user's behavior. The Situ theory does not fully meet the analysis of the intention of users with different identities in the social media environment. This paper's motivation is to analyze the user's intention sequence mode(s) in social media networks. The major contributions of this paper are two folds. One is to enrich and extend the Situ theory outreaching for social domain, that is the social media ecosystem, through newly and comprehensively considering user's changeable identity (including role and group), and the other is to propose a novel algorithm for users' behavior pattern analysis and mining. The important vision of the work is to further predict users' more and deeper intention and mental based on a large volume of previous actions.

The remaining parts of this paper are as follows: Section 2 shows the progress in related studies; the next section shows the extension of the Situ framework; Section 4 introduces the intention serialization algorithm; the experiment and its results of the serialization algorithm are in detail presented in Section 5; and finally conclusions are drawn.

2. RELATED WORKS

The significance and influence of the situation analysis theory and Situ framework on software engineering, as well as introduced the Situ framework in detail, which could provide users with personalized service by identifying the new intention of the user and the real-time update of service . Ming et al raised a spatial scenario analysis based on the Situ theory and the proposed (MR)2 paradigm promoted comprehensive decision-making which is conducive to the transformation of data, information, knowledge, and wisdom. Stated that, in a given environment, the user could share data with friends in the social circle through the part of the social service which they are involved in. So they put forward a SenseFacen framework to recommend services for users by using perceptual data from the user sensor network and multimedia information. Put forward an algorithm which considers the surrounding environment and social network relationship. This algorithm could make use of user's recognized situation, preference, and social network relationship to acquire user's nearest neighbours through the calculation of the user's comprehensive situation similarity, and predict the potential situation user preference to make a recommendation. The combined with the characteristics of Internet of things, to discuss information acquisition, modelling and intelligent processing etc by taking the situation awareness process as the main line. Hence, it becomes more and more important to employ a novel situational awareness for computing services to provide users with more personalized functions, including multimedia recommendation service, customized security and privacy one, and so forth.

The presented an improved N-gram prediction model to predict the possible future web access request of the user through the server log data . Bar-David et al stated that existing technology made an attempt to predict the location of moving user according to historical trajectory of moving objects, while ignoring the fact that the dynamic nature of the moving behavior may lead to errors in prediction. They proposed a type of context-aware position prediction algorithm based on various contexts to predict the future position of a vehicle. In order to allow smart phone users to access the service easily and timely, designed a recommendation mechanism to predict user's intention and activate appropriate service; an event-condition-behavior model and a rule induction algorithm was used to find out behavior patterns of smart phone users, and then, made use of their behavior pattern to predict and recommend the appropriate service for the users. In order to better understand users' intention in MSNs, we greatly need to explore users' online social behavior patterns.

Users' data are high noise and discrete in MSNs, especially mobile social networks , and these data cannot be used for analysis and mining in time. So, there is a need to collect and preprocess users' data before our next work.

The field of software engineering, not completely appropriate for the emerging application scenario of multimedia social networks. To sum up, in order to provide users with more personalized services in the multimedia social networks, this paper established a SocialSitu framework on the basis of Situ-analytics theory through comprehensively considering users' context and situation in MSNs. To obtain user's intention sequence, we proposed a novel algorithm for analyzing on SocialSitu(t) sequences of users through the improved GSP.

3. EXTENSION OF THE SITU FRAMEWORK IN MSNS

In MSNs environment, a large number of users may be in different groups with different roles. The roles of users in groups may cause them to generate different desires. Therefore, this paper extends and enriches Situ framework in social media, as defined below :

Definition 1 (Situation(t)): It represents the situation at t, which consists of a three-tuple, Situation(t)={d, A, E}. Where d refers to the desire of user at t; A refers to the action of the user which achieves the d; E refers to environmental context at t.

Definition 2 (SocialSitu(t)): It refers to the situation at t which is the extensional Situation(t) for the social domain. SocialSitu(t) is a four-tuple SocialSitu(t) = {ID ,d, A, E}. Here, ID refers to user's identity information; d refers to user's desire at t; A refers to user's behavior corresponding to d at the moment; E refers to environment information, including the terminal information which the user utilized.

Definition 3(ID): ID refers to the user's identity information; it is a two-tuple ID= {Group, Role}. In MSNs, there is a corresponding relationship between the user's role and group. When a user's role is changed, the user's behavior may also change.

Definition 4(Group): It refers to a small group formed in social media network because of a particular reason. It's a part of the whole social media network.

Definition 5 (Role, R): a user's role in MSNs. The role is a set R= {r1, r2, …, rn}, referring to RBAC96.

Definition 6 (Desire, D): It refers to what users want to achieve when using a social media service, namely, the user's purpose. It consists of a series of atom -desire (d), namely {d1, d2,..., dn} , d i (1 i n) refers to user's desire at i.

Definition 7 (Goal, G): the user's general target G= {g1, g2, … , gn} for MSNs.

Definition 8 (Intention, I): It refers to the SocialSitu(t) sequence of user from starting point to target achievement, namely I={SocialSitu(1), SocialSitu(2), ..., SocialSitu(n)} , n N , SocialSitu(1) refers to the starting point; SocialSitu(n) refers to the ending point when the target is achieved. Here, SocialSitu(t) sequence is directly correlated to the target achievement. Through the intention sequence, the user achieves the target, as shown in Fig.1.

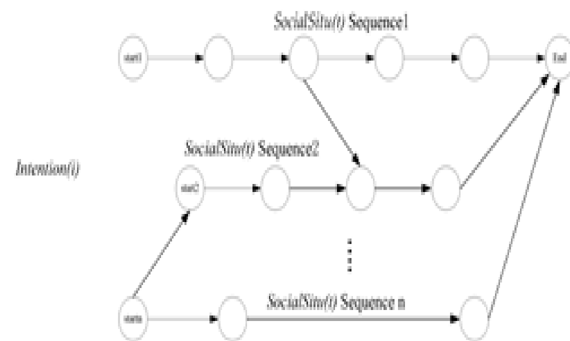

Fig.1. Intention sequence

In the figure, each point refers to SocialSitu(t) at a certain moment. The point startj (1 j n, j N) refers to the starting point of Intention(i). These starting points can be the same or different. End refers to the ending point of Intention(i). Each stripe of SocialSitu(t) sequence refers to the sequence composed by different SocialSitu(t) that the user passed from starting point to ending point. Except for the ending point, the same nodes may exist in each sequence of Intention(i). In the MSNs, there is at least one sequence which corresponds to the user's intention, namely i N , i 1 .

4. INTENTION SERIALIZATION ALGORITHM OF USER

All frequent SocialSitu(t) related to a certain goal achievement in a user's historical access record consist of an intention sequence. The user has at least one goal in MSNs, and this corresponds to at least one intention sequence. The user's intention sequence with a specific goal is saved to the database. The current sequence of a user is compared with intention sequences of the user in the database to predict the current intention of the user to make a rapid and timely response to the user's request and provide a personalized service, intention prediction flowchart is shown in Fig.2. A key problem in this paper is how to find out the user's Intention sequence.

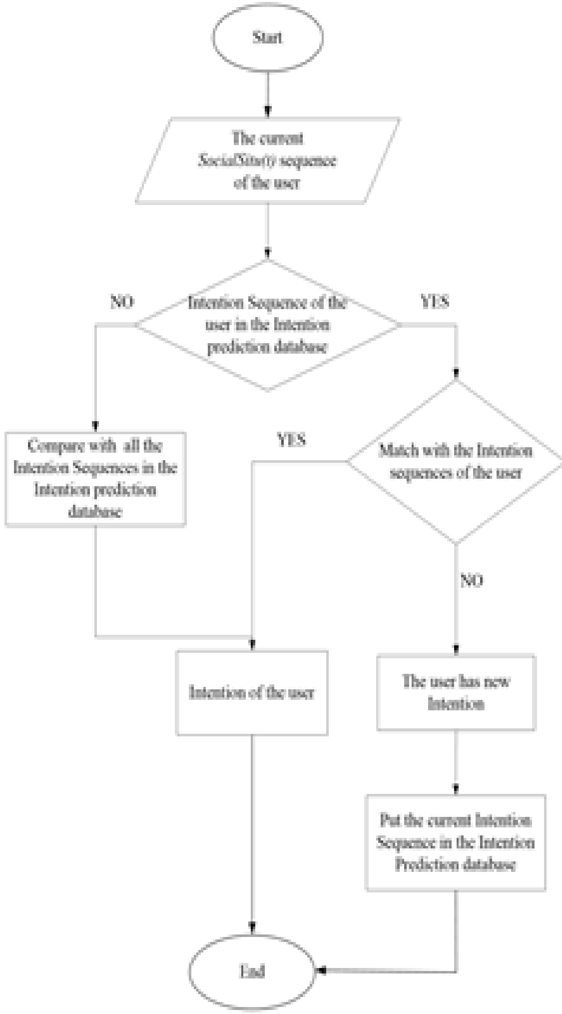

Fig.2. Intention prediction flowchart of the user

The association rule which was proposed by Agrawal et al in 1993 is used to find out the relationship among various items in a large quantity of data. DS is a set which represents the entire transaction set where each attribute is called as an item. The set including all items in a DS is named as the data item set, I= {i1, i2, ..., im}, |I|=m, m refers to the number of items in DS.

The association rule contains the following logic implication form: A B , wherein, A I , B I and A B ; item set A is the antecedent of the association rule; item set B is the result of the association rule; A B is the item set which corresponds to this rule.

Support: the number of item set R contained in the DS called as the supporting number of R, recorded as R.Support. The number of support forrule R S refers to number that the item sets R and S coexisting in DS. Therefore, the support for rule R S is Support (R S) = P (R S).

The item set satisfying the Min_Support is called the frequent item set. The rule satisfying the Min_Support and the minimum confidence threshold (Min_Conf) is the strong association rule. Therefore, Intention(i) serialization in this paper adopts the method based on the association rule to find out each sequence corresponding to the intention. The ending point of each Intention(i) sequence is used as the result of association rule, association rule is used to obtain the antecedent of association rule. Intention serialization algorithm is shown in Algorithm 1 and flowchart of intention serialization algorithm is shown in Fig.3.

The steps of serialization algorithm based on association rule are as follows:

(1) The web log database is scanned after data processing, the goal in definition 7 was identified in the database as the ending point of user in Intention(i),

recorded as G ' {g1' , g 2' ,…., gm'}, 1 m n , G '

(2) gi'obtained from Step (1) is used as a result of association rule. Each SocialSitu(t) is used as the antecedent of the association rule to calculate the Support of each rule, and find out the rule satisfying the Min_Support.

(3) The antecedents of the rule obtained from Step (2) are used to build a set L1, for set Lk in the length of k, where the link operation and pruning operation are used to generate a candidate sequence Ck+1 in the length of k+1. Then, scan data set DS, calculate the Support of each candidate sequence as the antecedent and g as result of

the association rule to generate sequence Lk+1 in the length of k+1, and Lk+1 is used as the seed set of the antecedent of new association rule.

(4) Step (3) is repeated until the new candidate sequence can no longer be generated, and all SocialSitu(t) sequences related the target g'i of Intention(i) is obtained.

(5) All SocialSitu(t) sequences corresponded to target g'i+1 are acquired and recorded as Intention(i+1). Then, Steps (2), (3), and (4) are repeated.

(6) Until there is no longer a new goal.

Link operation: if the sequence obtained after removing the first item of sequence pattern s1 is the same as the sequence obtained after removing the last item of sequence pattern s2, then s1 should be connected with s2. That is, the last item of s2 should be added into s1.

Pruning operation: if a certain sub-sequence of a candidate sequence pattern is not a sequence pattern, this candidate sequence pattern is unlikely to be a sequence pattern; therefore, it is deleted from the candidate sequence pattern.

Algorithm 1: INTENTION SERIALIZATION ALGORITHM BASED ON SITUATION-AWARE

Input: DataSet: DS, the Minimum Support: Min_Support, User’s Goal: G '

Output: SocialSitu(t) Sequence

Situ Behavior Analytics (DS, Min_Support, G')

```
Begin
for j1 to n    //n indicates the number of user’s goal
for t0 to T
Support( SocialSitu (t )g 'j ) = P ( SocialSitu (t ) g 'j );
endfor
if (Support ( SocialSitu (t )g 'j )>Min_Support)
L1  SocialSitu(t) ;    //the 1-frequent item sets L1
endif
for k2 to m and Lk 1   Null
Generate candidate sets Ck ;
Support ( Ckg 'j ) = P ( Ck    g 'j );
if (Support ( Ckg 'j )>Min_Support)
Lk   Ck ;
endif
endfor
Intention(i) = Lkg 'j ;
endfor
End
```

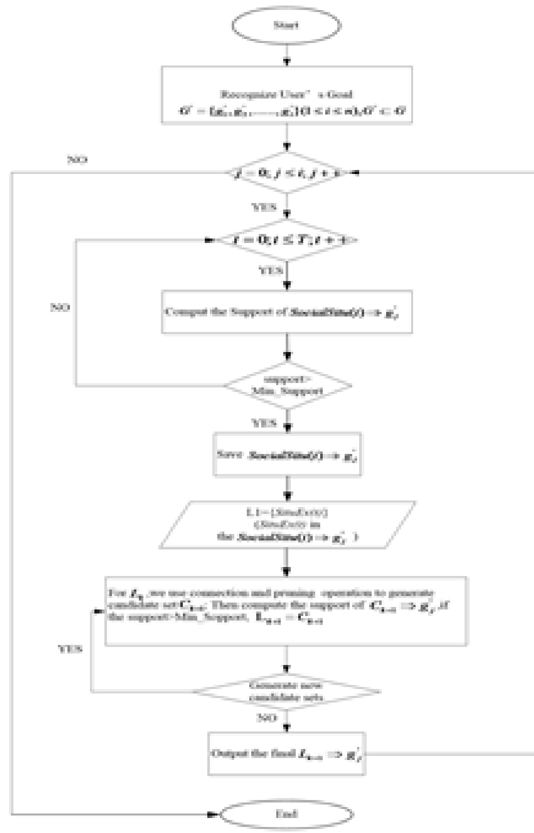

Fig.3. Flowchart of situational aware intention serialization algorithm

5. EXPERIMENT AND RESULT ANALYSIS

In the multimedia social network CyVOD [31] which is a prototype system we have achieved, supposing that users login and quit normally, the user first enters the goal before accessing the CyVOD, with the range of Goal. The user would quit the system when the goal is achieved. A complete conversation from logging in to quitting SocialSitu(t) sequence is tracked.

The four elements in SocialSitu(t) are acquired and enumerated below:

(1) ID: The user’s role and group are acquired in the database through the session information saved in the server. Users’ groups are the common registered user group and the advanced user group, which are corresponded with common users and VIP users, respectively.

(2) d: A user’s behavior in MSNs is an observable vector. However, a user’s desire is concealed. User’s behaviors are reflected in various states by a probability density distribution. For

example, when the user clicks into the login, the user's desire is corresponded with the login behavior access to the system.

(3) User's behavior A: In order to achieve d, the user's behavior may be an atomic action or a compound action, mainly referring to user's click and keyboard input behavior [16]. The user's behavior can be obtained through a web server log and the data change at a certain moment in the database.

In MSNs, the data in the web log are complex and the data preprocessing is required to transform these complex data into the data format required. The data preprocessing includes data cleaning, user identification, session identification, and data transformation, as shown in Fig.4, specific data preprocessing is shown as follows :

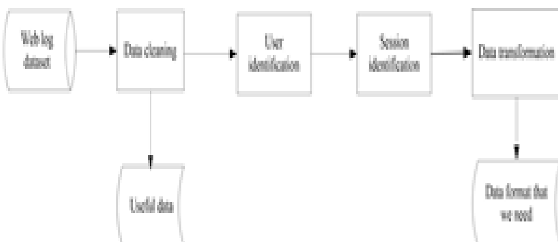

Fig.4. Data preprocessing for intention serialization algorithm

- Data cleaning: Irrelevant data record from the should be removed. For example: browsing errors, server errors, or client errors. These log data are insignificant to this research. The error information can be found in the status code in web log, and deleted.
- User identification: All data of the current user from a large quantity of logs is identified. This paper is aimed at registered users. Therefore, the user registered ID is employed to identify the user.
- Session identification: A session is a collection of pages accessed by a user during a certain period of time. A user's complete session is identified from logging in to quitting.
- Data transformation: Log data, which is continuous, is transformed into the data type required in this paper. These data are divided into discrete data points according to the time stamp, that is, transformed to be in the SocialSitu(t) four-tuple in definition 2.

The most common intentions are play and share. The two users' historical SocialSitu(t) data of play and share intentions are collected in this experiment, respectively. We collected hundreds of sessions' data to analysis the intention sequence patterns of User #1 with different IDs and User #2. There include a large volume of actions such as logining, searching, and so on, together with environmental information in sessions data, which are from the log data of CyVOD.net. By using the serialization algorithm in part 4, the basic sequence patterns of two users are obtained, as shown the intention sequence pattern of the Users.

Similarly, the intention sequence patterns of other users and final selection of Min_Support in MSNs can be concluded.

6. DATA COLLECTION

In order to acquire a large social media dataset while still maintaining user privacy, we developed an extraction tool deployed as a web service to collect anonymized data from volunteers' Facebook accounts. We recruited 1327 participants from across the U.S. that were at least 18 years old. Although we employed emailing colleagues and posting on Facebook for recruiting, most (>90%) of our participants were not affiliated with the companies nor were they friends with any of the involved researchers. Upon signing up for the study and giving informed consent, participants first answered an online survey. In that survey, we collected participants' Big Five Personality scores using a questionnaire adapted from International Personality Item Pool, as well as their demographic information (such as age, gender, marital status and education). After that, participants authorized social media applications to collect their social network activity data. In order to ensure an efficient data collection, we placed (generous) caps on the amount of each data type kept. For example, we restricted data to be from the last 365 days, and accumulated at most 1000 posts from each user. Like other standard Facebook applications, this application uses an open authorization standard to get permission from users, requiring neither passwords nor user names from participants. For privacy protection, study participants (and their Facebook friends) were assigned random unique IDs, and only these IDs are kept as an identifier in the extracted data set. In addition, we computed a set of high-level, aggregated statistics from each participant's activities. No raw text content or visual content from a user or the user's friends was collected. Instead, we replaced each distinct word occurring within text content with a unique integer. By these steps, we ensured that it is impossible to reconstruct personal identifiable information from users.

7. ACKNOWLEDGEMENTS

MSN and its mobile app prototypes, and also would like to thank the reviewers and editor for their valuable comments, questions, and suggestions. and Technology Grant , Natural Science Foundation of Henan Province Grant, and Project of the Cultivation Fund of Science and Technology Achievements of Henan University of Science and Technology. The authors gratefully acknowledge support for this work from DARPA through the SMIITE (Social Multimedia Investigation for Insider Threat Exposure) and GLAD-PC (Graph Learning for Anomaly Detection using Psychological Context) projects funded as part of the ADAMS (Anomaly Detection At Multiple Scales) program. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the government funding agencies.

8. CONCLUSIONS

The existing MSNs environment increasingly requires situation awareness. Users' environment and behavior are dynamic, and an individual's intention is also to change. In order to adapt to the dynamic changes of user identities in the social domain, this paper extends and enriches the Situ theory, and builds a SocialSitu framework for the social media networks. We design and achieve the intention serialization algorithm in multimedia social networks. The user's frequent intention sequence mode is obtained through the intention serialization algorithm. When the user's identify changes, we conclude his behavior pattern with different ID, and prove that different SocislSitu(t) sequences are acquired in the same Min_Support with the same intention when his role and group change. In the future works, the existing intention sequence patterns of the user could be adopted to predict the user's more and deeper intentions. Besides, we will employ the SocialSitu and the proposed algorithm to improve multimedia recommendation system and some killer applications in MSNs.

REFERENCES

[1] B. De Meester, R. Verborgh, P. Pauwels, W. De Neve, E. Mannens, and R. Van de Walle, "Towards robust and reliable multimedia analysis through semantic integration of services," Multimedia Tools Appl., vol. 75, no. 22, pp. 14019-14038, Nov. 2016.

[2] Z. Zhang and B. B. Gupta, "Social Media Trustworthiness and Security: Overview and New Direction,"Future Generation ComputerSystems, submittedforpublication.

[3] B. N. Schilit and M. M. Theimer, "Disseminating Active Map Information to Mobile Hosts," IEEE Network, vol. 8, no. 5, pp. 22-32,Sep/Oct1994.

[4] K. Srinivasan, P. Agrawal, R. Arya, N. Akhtar, D. Pengoria, and T. A. Gonsalves, "Context-aware, QoE-driven adaptation of multimedia services," 5th International Conference on Mobile Wireless Middleware, Operating Systems, and Applications, pp. 236-249,Nov.2012,doi: 10.1007/978-3-642-36660-4_17.

[5] Radhakrishna,S.A.Aljawarneh,P.V.Aljawarneh,a nd K.-K. R. Choo, "A novel fuzzy gaussian-based dissimilarity measure for discovering similarity temporal association patterns," Soft Comput., preprint, 18 Nov. 2016, doi: 10.1007/s00500-016-2445y.

[6] D. Fisher, M. Smith, and H. Welser, "You are who you talk to," in HICSS'06, 2006, vol. 3, p. 59b.

[7] F. Benevenuto, T. Rodriguesy, M. Cha, and V. Almida, "Characterizing User Behavior in Online Social Networks," in IMC'09, 2009, pp. 49–62.

[8] M. Maia, J. Almeida, and V. Almeida, "Identifying User Behavior in Online Social Networks," in SocialNets '08, 2008, pp.

[9] J. Brenner, "Pew Internet: Social Networking," Pew Internet, 2013. [Online]. Available: http://bit.ly/15rpKKY. [Accessed: 07-May-2013].