

Assessing Post Usage for Measuring the Quality of Forum Posts

Kevin Chai, Pedram Hayati, Vidyasagar Potdar, Chen Wu and Alex Talevski
Digital Ecosystems and Business Intelligence Institute
Curtin University, Perth, Australia
{kevin.chai, pedram.hayati}@postgrad.curtin.edu.au
{vidyasagar.potdar, chen.wu, alex.talevski}@cbs.curtin.edu.au

Abstract—It has become difficult to discover quality content within forums websites due to the increasing amount of User Generated Content (UGC) on the Web. Many existing websites have relied on their users to explicitly rate content quality. The main problem with this approach is that the majority of content often receives insufficient rating. Current automated content rating solutions have evaluated linguistic features of UGC but are less effective for different types of online communities. We propose a novel approach that assesses post usage to measure the quality of forum posts. Post usage can be viewed as implicit user ratings derived from their usage behaviour. The proposed model is validated against an operational forum using Matthews Correlation Coefficient to measure performance. Our model serves as a basis of exploring content usage to measure content quality in forums and other Web 2.0 platforms.

Keywords: forums, content usage, content quality assessment, user generated content, web usage mining

I. INTRODUCTION

Forums are web bulletin boards that facilitate discussions and can host large amounts of User Generated Content (UGC). Due to their popularity, there are millions of operational forums on the Web. Users are finding it more difficult to discover quality content within forums as well as other Web 2.0 websites in a timely manner which is referred to as information overload.

This is a compounding problem with the rapid increase of content on the Web. [1] reports that in 2006 there was 161 billion GB of data on the Web with UGC representing the largest portion of this data. Many Web 2.0 websites rely on users to manually and explicitly rate the quality of content to handle this issue [2].

However there are inherent problems in relying solely on user ratings. The main problems are that:

A large percentage of content receives a lack of user rating [3]

- Ratings are first needed before people can be made aware of content quality (known as the ramp-up problem) [4]
- Assumes that user feedback is honest and users have sufficient knowledge to provide meaningful ratings [5-8]

- Premature negative consent below a visibility threshold no longer presents a post for further viewing and rating [3, 9]
- Content that is rated incorrectly is often not rectified by moderators [10]
- Reliance on explicit user ratings results in an ongoing problem if the amount of UGC is created at a faster speed than which ratings are generated.

The main purpose of this paper is to propose a novel and automated approach to measuring the quality of forum posts. More specifically, the contributions of this paper are to:

- Present a method of measuring post usage and evaluate its feasibility in measuring the quality of forum posts.
- Measure the quality of posts in an operational forum and compare these results to labelled data using a supervised machine learning approach.

The specific problems of automatically measuring the quality of UGC in forums will now be discussed.

II. PROBLEM

A number of frameworks have been developed to automatically measure the quality of UGC [2]. However, many of these frameworks have relied on content analysis which is often computationally complex and produces a language dependent framework (e.g. only works for content in English) [2]. Examples would include evaluating content based upon its punctuation, capitalisation, spelling errors, readability etc... [3, 10-12].

However, the characteristics of forum posts indicate that traditional content analysis techniques may not be readily applicable to forum communities because:

- A single post is generally constructed as text fragments relating to other posts [3, 10]
- A single post is often incomplete, error-prone and poorly structured [13]
- Different forum communities have distinct terminology and public forums generally do not follow proper linguistic rules or a formal style of writing [3, 10]

Therefore we propose a novel approach that evaluates post usage to predict the quality of forum posts. Post usage is considered as implicit user ratings derived from monitoring their content usage behaviour. The problem of measuring post quality is formalised in following subsection and our approach is detailed in Section III.

A. Maintaining the Integrity of the Specifications

The problem of measuring post quality in discussion forums is formally defined as a multi-class classification problem. The forum data set D is described by a set of posts $P = \{p1, p2, \dots, pn\}$ where n denotes the total number of posts and a set of post usage features $F = \{f1, f2, \dots, fm\}$ where m denotes the number of features. The data set also contains a target attribute C that is called the post quality class. The class attribute C has a set of discrete values, $C = \{c1, c2, c3\}$ which describes posts P that are categorised in low ($c1$), medium ($c2$) and high ($c3$) quality classes.

III. CONCEPTUAL MODEL

The main philosophy behind our model is that, “actions speak louder than words”. More specifically, rather than asking users to explicitly rank the quality of forum posts, we monitor behaviour of users in consuming forum posts (e.g. time users spend reading a post). By proposing this method, we are able to address problems described in Section I with the exception of the ramp-up problem. We are unable to resolve this problem because time is first required to track content usage before post quality can be predicted. The fundamental assumption of this model is that community usage of posts provides insights into its quality.

Mouse tracking techniques is used to measure user consumption of forum posts. Mouse tracking has largely been used in software usability studies, which are discussed in Section VI.B. To the best of our knowledge no existing work has been dedicated to evaluating the usage of forum posts. A graphical representation of the conceptual model is presented in Figure 1 and each component is discussed in detail throughout the paper.

A. Post Usage Tracking

A module is developed to track post usage that consists of 4 components:

1. Page navigation: tracks the URL, date and time, user’s session, user’s IP address, user’s referrer URL and the user’s agent.
2. Post view: tracks the post, date and time, user’s session and action (mouse enters or leaves a post).
3. Post selection: tracks the post, date and time, user’s session and the text that is selected.
4. Post copy: tracks the post, date and time, user’s session and the text that is copied.

Traditionally, user navigation tracking is done through processing web server logs but this method is unable to link a specific page request to a specific forum user account and their session [14, 15]. Therefore, our model tracks page navigation

within the forum through web scripting to handle user and session identification. Page navigation tracking is required to assist in calculating post dwell time (see Section IV.A.3).

We adopt standard web technologies and techniques to ensure our approach is applicable for operational forums. Asynchronous JavaScript and XML (AJAX) is used to track post views by recording when a user’s mouse enters and leaves a post. Text selection is recorded each time text is selected and the mouse button is released within a post. Additionally, post copy is recorded when a user presses a keyboard shortcut to copy selected text within a post. Our module has a small footprint and did not degrade the performance of the forum during testing.

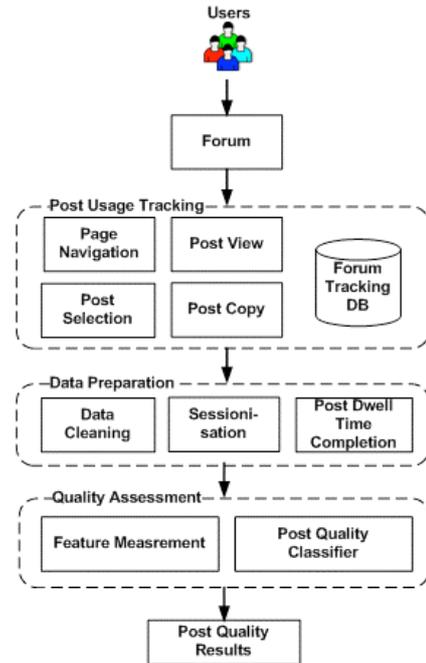


Figure 1. Conceptual Model

B. Quality Assessment

We propose the use of a supervised machine learning technique known as Support Vector Machines (SVM) to measure forum post quality. We propose the evaluation of four post usage feature categories including post view counts, the time users spend viewing a post, mouse and keyboard interactions in our model as presented in Table I.

Firstly, post view counts are evaluated to assess its amount of use and popularity. Secondly, the amount of time users dwell within posts provides us with an indicator of reading time. Lastly, mouse and keyboard interactions are captured through text selection and copies. Text selection is evaluated to provide evidence that a user is active within a post. For example, users may select keywords or highlight text during reading out of habit. Text copies are captured to measure the utility of text. i.e. the user is planning on pasting that text somewhere else.

Additionally, posts which are older and longer in terms of text have more opportunity to accumulate more dwell time, text selection and text copies. Therefore, post age and number of words within a post are incorporated as features in our model.

We use Matthews Correlation Coefficient (MCC) to measure the performance of our model. MCC is considered one of the best methods for evaluating classifier prediction on the imbalanced data [16]. Additionally, MCC provides a better measure of performance than the evaluation of the confusion matrix of a classifier [17].

IV. EXPERIMENT

A. Data Collection

Data has been collected from an online gaming forum from the 21st of July 2009 to the 16th of October for a total duration of 88 days. A number of data preparation tasks are applied to the dataset prior to experimentation, which are now discussed.

TABLE I. POST USAGE FEATURES

| Code | Name |
|------|--|
| F1 | Total # of views |
| F2 | Total # of distinct user views |
| F3 | Total dwell time |
| F4 | Average dwell time |
| F5 | Total # of text selections |
| F6 | Ratio # of text selections by total # of views |
| F7 | Ratio # of text selections by total dwell time |
| F8 | Ratio # of text selections by average dwell time |
| F9 | Total # of characters selected in text |
| F10 | Average # of characters selected in text |
| F11 | Total # of text copies |
| F12 | Ratio # of text copies by total # of views |
| F13 | Ratio # of text copies by total dwell time |
| F14 | Ratio # of text copies by average dwell time |
| F15 | Total # of characters copied in text |
| F16 | Average # of characters copied in text |
| F17 | Age in days |
| F18 | Number of words |

1) Data Cleaning

1. Removal of monitoring data that includes data collected from researchers monitoring the forum and therefore unintentionally contributing to the post views.
2. Removal of post view records with less than 1 second of dwell time. Users can scroll down a page and have their mouse enter and leave posts in quick succession

resulting in these post views being recorded. We believe that the majority of these records with less than 1 second of dwell time are navigation instances and not of users reading the post.

3. Removal of duplicate post selection and copies. We found records of users selecting and copying the same text on numerous instances within a short period of time possibly due to their browsing habits. Duplicate post selection and copy records are removed.
4. Removal of crawler robot (e.g. Google Bot) and anonymous user records. The majority of forum posts are only viewable by registered forum members so including anonymous user post view records would result in collecting more post usage for public posts when compared with member only posts

2) Sessionisation

Users within this forum are able to use cookie-based logins (i.e. log me in automatically feature) to maintain the same session until cookie expiry. A common heuristic used in web usage mining is to cease a user session after 30 minutes of inactivity [18] and this is adopted in our experiment.

3) Post Dwell Time Completion

There are instances where the post view dwell time can not be calculated directly from when the mouse enters and leaves a post. Examples of these scenarios include:

1. When the forum thread is loaded and the mouse is already inside the post. Therefore a mouse enter event is not registered but a mouse leave event is recorded when it leaves the initial post. We query the page navigation records for the time the page is loaded to calculate the dwell time.
2. A user has left the page without their mouse leaving a particular post (e.g. clicking on a hyperlink within a post). The page navigation records are queried for the time the user has navigated to another page. If this record does not exist (i.e. when the user closes their web browser or a new session was started) then we adopt a commonly used web usage technique for time spent on viewing the last page in a user session. i.e. we calculate the average dwell time of all posts within the page and assign that average value to the last post visited.

B. Dataset

The dataset after preparation contains 54 users, 114 topics and 532 posts. The dataset is split into 2 subsets with two thirds of used for training and the remainder used for testing. The rationale behind splitting the dataset is to treat the test set as future posts that require post quality prediction.

Two forum administrators were asked to rate the quality of all posts (1 to 5 star post ratings). The Pearson correlation is 0.515, Kendall's tau is 0.428 and Spearman's rho is 0.499 with sig values less than 0.01 when measuring the inter-rater agreement. These results show there is certainly a moderate

and positive correlation (agreement) between the administrators' post ratings.

We further group posts into three quality classes due to the imbalanced distribution of posts belonging to each quality class. There are 9 possible average post rating scores and these are distributed equally to each quality class. More specifically, $low = \{1, 1.5, 2\}$, $medium = \{2.5, 3, 3.5\}$ and $high = \{4, 4.5, 5\}$. Details of the training and test dataset are presented in Table II.

TABLE II. TRAINING AND TEST DATASETS

| Quality Label | Training | Test | Total |
|---------------|------------|------------|-------------------|
| Low | 196 | 93 | 289 (54%) |
| Medium | 105 | 61 | 166 (31%) |
| High | 53 | 24 | 77 (14%) |
| Total | 354 | 178 | 532 (100%) |

C. Results

A supervised machine learning technique known as Support Vector Machines (SVM) is trained and used to predict post quality on the test dataset. We utilise a popular SVM library known as LibSVM [19] with a Radial Basis Function kernel, feature scaling and 5-fold cross validation on the training and test dataset. A classifier accuracy 65.17% was achieved from our experiment with the kernel parameters $C=512$ and $\gamma=0.03125$ in addition to a cross-validation rate of 67.79%.

A confusion matrix of the resulting classifier is presented in Table III, which shows the number of true positives, false positives, true negatives and false negatives for each quality class. This matrix indicates that the classifier is quite accurate in predicting low quality posts achieving roughly 96% accuracy (90/93). However the strength of the classifier degrades in predicting higher quality classes as the accuracy of predicting medium quality posts is around 36% (22/61) and high quality posts is 16% (4/24).

TABLE III. CONFUSION MATRIX

| | | Actual | | |
|-----------|--------|--------|--------|------|
| | | Low | Medium | High |
| Predicted | Low | 90 | 35 | 8 |
| | Medium | 2 | 22 | 12 |
| | High | 1 | 4 | 4 |

In order to provide another metric for classifier evaluation we calculate the Matthews correlation coefficient (MCC) for each of the three post quality classes. This metric provides a correlation value between [-1,1] and where -1 represents inverse prediction, 0 represents random prediction and 1 represents perfect prediction.

Firstly, a MCC_{low} value of 0.5308 was achieved which represents a moderately strong ability of our classifier to predict low quality posts. Secondly, a MCC_{medium} value of

0.2885 was calculated which shows a weak to moderate ability for our classifier to predict medium quality posts. Lastly, MCC_{high} achieved a value of 0.2092, which shows a weak to moderate ability of classifier to predict high quality posts. These MCC results indicate that the performance of our classifier degrades in classifying posts of higher quality.

V. DISCUSSION

The post usage classifier achieved high accuracy in classifying low quality posts. However, the classifier found difficulty in separating medium quality posts with low quality posts and high quality posts with medium posts based upon their usage. This highlights that possibility that medium and high quality posts have a larger and more diverse range of usage behaviour. Additionally, there are a number of reasons why this may have occurred.

Firstly, the post usage tracking module can be enhanced in the future to capture more usage data. Users have been known to use the scrollbar when viewing web content rather than hovering their mouse within a post [20, 21]. Additionally users have been profiled in existing research to leave their mouse in blank space to the side rather than over content [22]. Finally, users can read multiple posts on a page but may not necessarily hover their mouse within each post while reading. Our system is unable to capture post usage in these scenarios that may or may not have impacted on the performance of classifying medium and high quality posts.

Secondly, the purpose of this study was to evaluate the feasibility of assessing post usage to infer post quality. Popularity and relevancy (i.e. the dimensions we are measuring through post usage) are only two dimensions of content quality [2]. A strategy to improve the classifier for medium and high quality posts would be to gather more evidence by evaluating other quality dimensions through content, meta-content, temporal and structural features of posts, which we will incorporate in future, work.

Thirdly, content visibility is not considered in our model. Posts that are more visible (e.g. a post displayed on the first page of a thread) is likely to accumulate more usage than content with less visibility (e.g. a post displayed on the third page of a thread out of a 6 page thread). This can also be the result of forum security where only moderators and administrators are allowed to view specific threads which would therefore accumulate less community usage. This model can be improved by normalising post usage data based upon post visibility within a thread to measure quality.

This observation suggests the possibility to evaluate post quality within the context of a specific thread rather than across different threads. For example, forum functionality could be developed for users to order posts by their usage and quality within a thread. As previously discussed, posts are generally related to other posts and future research can explore how to link related posts and to measure the quality sub-thread discussions.

While post quality classification can assist users in browsing, it becomes less meaningful when the number of posts within each quality class becomes large. Posts were

grouped into low, medium and high quality classes for this experiment. For example, if there are 2,000 high quality posts then without some user input (i.e. search query), it becomes difficult for a user to browse through all of these posts. Additionally, providing post quality information as a number in one dimension may not be as useful to users. Therefore, a more difficult and challenging avenue of improving this model is to generate post quality rankings and/or assessing the quality of posts over number of quality dimensions.

Lastly, user fraud is not handled in our model. Fraud for our model is similar to search engine optimisation in which users can exploit search ranking algorithms to position their website higher in search results. In the forum domain, users may wish to improve the quality ranking of their posts to increase readership and possibly provide links to their websites through their forum signature. Additionally, malicious users may wish to degrade the performance of the post quality model to reduce a specific forum's usability.

VI. RELATED WORK

A. Forum Post Quality

An extensive review was conducted by [2] which evaluated 19 content quality related assessment frameworks for various Web 2.0 platforms including wikis, forums, question & answering portals and weblogs. A number of content quality dimensions were identified from the surveyed frameworks. Our post usage model has been designed to measure the implicit user feedback, relevancy and popularity dimensions of quality.

[3] was first to propose a model for specifically measuring the quality of forum posts and classified posts into two quality classes (high and low) by assessing surface, lexical, syntactic, similarity and forum specific post features. [10] extended the scope of this research by classifying posts into 3 quality classes (low, medium and high) and evaluated features such as relevance, originality, post component, surface and forum-specific features. [23] employed link analysis techniques in ranking postings in terms of their importance to improve search engine rankings of forum posts.

[11] evaluated usage statistics of questions and answers in a question and answering (Q&A) portal (Yahoo! Answers) to find high quality content. Additionally, the number of times an answer was copied by users was proposed as a feature by [24] for measuring the quality of answers in Naver (Korean Q&A portal).

We adopt the idea of measuring content usage in Q&A portals to forums. To the best of our knowledge, no work has been conducted to measure post quality by evaluating the post usage behaviour of the forum community. Additionally, [25] proposed that age and visibility should be considered as factors in measuring cooperation and quality of Wikipedia articles. We incorporate post age as a feature for normalisation as older posts have more opportunity to accumulate more usage than younger posts.

B. Eye to Mouse Movement

Existing research has discovered a positive correlation between eye and mouse position / movement on a computer screen [26, 27]. Additional work has identified that the eye to mouse relationship is also apparent for website usage [22, 28, 29].

These studies provide evidence that to a certain extent, mouse tracking can serve as an alternative eye tracker. The main advantage of using mouse tracking over traditional eye tracking equipment is that it can be used to track all users, it is inexpensive and requires fewer resources in comparison (e.g. laboratory time and for participants to visit the labs etc...) [29].

As a result, later research has employed mouse tracking for evaluating usability of web portals, e-learning systems, web library catalogues and web search results [20, 30, 31]. [22] discovered 3 types of eye-mouse movement patterns for consuming web search results which include:

1. Keeping the mouse still and away from the place they are reading (e.g. in whitespace to the right or left of content).
2. Using the mouse as a reading aid to keep their place on the page while reading.
3. Using the mouse to mark an interesting result while they continue to check other results.

Our proposed post usage tracking component is able to capture eye-movement patterns 2 and 3 but it is unable to capture pattern 1. Additionally, work conducted by [32] discovered that the time spent on a webpage and the amount of scrolling with the mouse and keyboard were good implicit indicators of user interest. These features were measured through a custom-built web browser.

While the use of mouse tracking to calculate content dwell time is not new, its application to quality assessment of forum content and more specifically for each forum post is novel. The implementation of our post usage tracking component is most similar to [30]. The advantage of this approach is its simplicity and practicality in being applied to any forum or Web 2.0 website. However, we are unable to capture the amount of scrolling as evaluated in [32] due to our implementation.

VII. CONCLUSION

We have presented a novel model that evaluates post usage to measure post quality in this paper. This model has been validated against an operational forum and we used MCC to measure the performance of our proposed model. Our model achieved promising results in identifying low quality posts but requires improvement in classifying medium and high quality posts. From our results we discussed ways our post quality prediction model can be improved through assessing more quality dimensions, improving our tracking module and normalising post usage by the visibility of a post within a thread. Additionally, we plan on integrating content features with our usage model to possibly improve its performance.

Future research is being dedicated to tracking keyboard and web form usage behaviour in addition to incorporating our

model in assessing the contribution of content providers. Other areas of research could evaluate the post usage model performance in assisting search result ranking and forum user browsing as well as its application to other types of Web 2.0 platforms such as weblogs, wikis and Q&A portals for content quality measurement. Additionally, future research could be directed to understanding the psychology of how people interact and contribute content within a forum environment.

ACKNOWLEDGMENT

Kevin Chai thanks the forum administrators, Kieran Boyce and Daniel Clark who participated in this experiment by providing post quality ratings.

REFERENCES

- [1] M. Brodie, "The End of the Computing Era: Hephaestus meets The Olympians," Presented in IEEE/IES Digital Ecosystem and Business Intelligence Conference. Retrieved June 16, 2008, from http://www.michaelbrodie.com/documents/Digital_Ecosystems_Brodie_Ecosystems_2008_V2.pdf, 2008.
- [2] K. Chai, V. Potdar, and T. Dillon, "Content Quality Assessment Related Frameworks for Social Media," *Lecture Notes of Computer Science (LCNS) and Proceedings of the International Conference on Computational Science and Its Applications – ICCSA*, pp. 800-814, 2009.
- [3] M. Weimer and I. Gurevych, "Predicting the Perceived Quality of Web Forum Posts," *Proceedings of the Conference on Recent Advances in Natural Language Processing*, 2007.
- [4] P. Dondio and S. Barrett, "Computational Trust in Web Content Quality: A Comparative Evaluation on the Wikipedia Project," *Informatica*, vol. 31, pp. 151-160, 2007.
- [5] J. Han and Y. Liu, "Dubious Feedback: Fair or Not?," *Proceedings of the First International Conference on Scalable Information Systems*, vol. 152, 2006.
- [6] H. W. Lauw, E. P. Lim, and K. Wang, "Summarizing review scores of "unequal" reviewers," *Proceedings of the SIAM Conference on Data Mining (SDM07)*, pp. 539-545, 2007.
- [7] E. P. Lim, B. Q. Vuong, H. W. Lauw, and S. A., "Measuring Qualities of Articles Contributed by Online Communities," *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 81-87, 2007.
- [8] M. E. Nussbaum, K. Hartley, G. M. Sinatra, R. E. Reynolds, and L. D. Bendix, "Enhancing the Quality of On-Line Discussions," *Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA*, 2002.
- [9] C. Lampe and P. Resnick, "Slash(dot) and Burn: Distributed Moderation in a Large Online Conversation Space," *Proceedings of ACM CHI 2004 Conference on Human Factors in Computing Systems*, pp. 543-550, 2004.
- [10] N. Wanas, M. El-Saban, H. Ashour, and W. Ammar, "Automatic Scoring of Online Discussion Posts," *Proceeding of the 2nd ACM Workshop on Information Credibility on the Web*, pp. 19-26, 2008.
- [11] E. Agichtein, C. Castillo, D. Donato, A. Gionis, and G. Mishne, "Finding High-Quality Content in Social Media," *Proceedings of the international conference on Web search and web data mining (WSDM08)*, pp. 183-194, 2008.
- [12] J. E. Blumenstock, "Size Matters: Word Count as a Measure of Quality on Wikipedia," *Proceedings of the 17th international conference on World Wide Web*, pp. 1095-1096, 2008.
- [13] A. K. F. Lui, S. C. Li, and S. O. Choy, "An Evaluation of Automatic Text Categorization in Online Discussion Analysis," *Proceedings of the Seventh IEEE International Conference on Advanced Learning Technologies (ICALT)*, pp. 205-209, 2007.
- [14] J. Borges and M. Levene, "Data Mining of User Navigation Patterns," *Lecture Notes in Computer Science*, vol. 1836, pp. 92-111, 1999.
- [15] R. Cooley, B. Mobasher, and J. Srivastava, "Data Preparation for Mining World Wide Web Browsing Patterns," *Knowledge and Information Systems Journal*, vol. 1, pp. 5-32, 1999.
- [16] P. Baldi, S. Brunak, Y. Chauvin, C. A. F. Andersen, and H. Nielsen, "Assessing the Accuracy of Prediction Algorithms for Classification: An Overview," *Bioinformatics*, vol. 16, pp. 412-424, 2000.
- [17] B. W. Matthews, "Comparison of the Predicted and Observed Secondary Structure of T4 Phage Lysozyme," *Biochim Biophys Acta*, vol. 405, pp. 442-451, 1975.
- [18] M. Spiliopoulou, B. Mobasher, B. Berendt, and M. Nakagawa, "A Framework for the Evaluation of Session Reconstruction Heuristics in Web-Usage Analysis," *INFORMS Journal on Computing*, vol. 15, pp. 171-190, 2003.
- [19] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," *Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>*, 2001.
- [20] D. S. L. Chee and S. G. Khoo, "Users' Mouse/Cursor Movements in Two Web-Based Library Catalog Interfaces," *Proceedings of the International Conference on Work with Computing Systems (WWCS)*, vol. 640-645, 2004.
- [21] J. Goecks and J. Shavlik, "Learning Users' Interests by Unobtrusively Observing Their Normal Behavior," *Proceedings of the 5th international conference on Intelligent user interfaces*, pp. 129-132, 2000.
- [22] K. Rodden and X. Fu, "Exploring How Mouse Movements Relate to Eye Movements on Web Search Results Pages," *Web Information Seeking and Interaction Workshop*, 2007.
- [23] Z. Chen, L. Zhang, and W. Wang, "PostingRank: Bringing Order to Web Forum Postings," *Lecture Notes in Computer Science*, vol. 4993, pp. 377-384, 2008.
- [24] J. Jeon, W. B. Croft, J. H. Lee, and S. Park, "A Framework to Predict the Quality of Answers with Non-Textual Features," *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 228-235, 2006.
- [25] D. M. Wilkinson and B. A. Huberman, "Cooperation and quality in wikipedia," *Proceedings of the 2007 international symposium on Wikis*, pp. 157-164, 2007.
- [26] M. C. Chen, J. R. Anderson, and M. H. Sohn, "What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing," *Conference on Human Factors in Computing Systems*, pp. 281-282, 2001.
- [27] F. Mueller and A. Lockerd, "Cheese: Tracking Mouse Movement Activity on Websites, a Tool for User Modelling," *Conference on Human Factors in Computing Systems*, pp. 279-280, 2001.
- [28] A. L. Cox and M. M. Silva, "The Role of Mouse Movements in Interactive Search," *Proceedings of the 28th Annual Meeting of the Cognitive Science Society*, 2006.
- [29] L. Cooke, "Is the Mouse a "Poor Man's Eye Tracker"?", *Annual Conference Society For Technical Communication*, vol. 53, p. 252, 2006.
- [30] C. Ullrich and E. Melis, "The Poor Man's Eyetracker Tool of ActiveMath," *Proceedings of the World Conference on E-Learning in Corporate Government Healthcare and Higher Education eLearn*, pp. 2313-2316, 2002.
- [31] G. Costagliola, F. Ferrucci, V. Fucella, and L. Zurolo, "LOG4P: An Innovative Logger Framework for Web Portals," *Proceedings of the 3rd International Conference on Web Information Systems and Technologies*, vol. 2, 2007.
- [32] M. Claypool, P. Le, M. Wased, and D. Brown, "Implicit interest indicators" *Proceedings of the 6th international Conference on Intelligent User Interfaces*, vol. pp. 33-40, 2001.