Adaptive Feedback System for Websites based on User Classification

Garric Mathias Department of Computer Engineering St. Francis Institute of Technology Naina Bharadwaj Department of Information Technology St. Francis Institute of Technology Falguni Bharadwaj School of Engineering and Applied Sciences State University of New York, Buffalo

ABSTRACT

In this paper, a feedback system for websites has been proposed which adapts according to user's familiarity of the website thus helping in understanding areas of improvement from those users who use these features most, using the example of an e-commerce website. The feedback of users is analyzed by classifying them into discrete levels from 1 to 5 based on their interaction with the website and the Pareto Principle is used to reason this classification. Every questionnaire is based on the complexity of the level in which the user has been classified. An algorithm has been proposed for user classification based on factors such as how many services the user has availed and how active he is on the website. This system is adaptive in nature as the feedback form adapts itself to the user levels thus extracting relevant information.

General Terms

Data Mining, Data Analytics, Data Collection

Keywords

User Feedback, Adaptive feedback, Website Improvement, Pareto Principle

1. INTRODUCTION

Users can prove to be helpful in identifying problems as well as providing insights about solutions for different functionalities of a website; hence user feedback is essential for the efficient working of a website [1]. Feedbacks obtained by monitoring user behavior do not cover all the information required for website or service improvement [2]. There exist certain features that need explicit feedback from the user as highlighted in this paper [4]. This proposed system aims to obtain feedback explicitly from the user for different site functionalities. An adaptive approach has been used for better data collection based on user history and behavior. Existing websites need to obtain feedbacks from a large number of users to identify the existing problems and their corresponding user suggested solutions; thus there is a need for a structured and systematic approach that extracts the best out of user experience.

In this paper, a system that provides structured modules for adaptive feedback forms according to the user classification based on their usage of the website has been proposed. The user classification is done based on the Pareto Principle [5]. The example of an e-commerce website has been considered. It needs direct user feedback to identify limitations as well as improvements of various sectors. These range from usability of the website to the quality, delivery and maintenance of the equipment sold.

2. LITERATURE REVIEW

Considerable research has been carried out in the areas of implicit user feedback to improve user experience on websites. In [1], Gabriella Pasi talks about monitoring user's actions implicitly and inferring their predilections from those. While implicit feedback may give an insight about a user's preferences, it does not expand on what improvements could be made to the websites to draw more users. In [2], User interest modeling is used to provide personal information services which uses explicit feedback from the users. Yi Zhu et al have used user's clicking behavior, vote behavior, bookmark behavior and time-attenuation to update the working of the model. The explicit feedback techniques that are used in [2] are not explicit enough as it does not necessarily give an idea on what features will result in improvement of the website. [3] gives a better understanding on improving the development of websites by providing a multidimensional approach which evaluates user satisfaction. Farid Aulia Tanjung et al propose surveys as a method of evaluation based on performance metric needs. Although a good approach, it considers every user as equal which does not reflect the real world. Gerson Tontini in [4] provides attributes on which to quiz the users to improve customer satisfaction but doesn't take into account the diversity of user interaction with the websites. One user might use one set of features more than another set and hence might be able to give more insight on those features than the other users. This system makes use of both implicit and explicit feedback to identify the level of the user, starting from a newbie to an experienced one. The proposed system gives a questionnaire appropriate to the user based on his level which is calculated on the basis of what services he has used and for how long. In [5], Qing Li proposes using Pareto principle for improving ecommerce of eco bags. Their research is highly lacking actual factors and are designed for a very limited system. The survey has been combined taking approach with the Pareto Principle to optimize the adaptive system such that the focus is on 20% of the top user's feedback since they are the ones that use the system most. In [6], new methods of sentiment analysis have been mentioned that are used in extracting opinions out of feedbacks. In [7], Dilip Kumar et al use user's local instance repository and world knowledge base to discover background knowledge about user and create personalized ontology. This is a disadvantage as entry table which acts as WKB is essentially the same for most users. Also, user's local instance repository would not provide apt classification level of user as user himself won't be able to classify himself according to the website's standards. Madhuka Udantha in [8] recognizes website user access patterns some are common but do not elaborate on how to make use of the rare user access patterns and how user patterns can be used to form clusters. Although Dora Dzvonyar et al in [9], integrated the feedback into the

application, it was limited in generalizability of results and its validity was influenced by the constricted setting environment.

3. PROPOSED WORK

Our proposed work includes three basic modules for adaptive feedback collection as seen in Figure 1.





3.1. Defining Basic Sections

In this module, the various sections aimed for data collection are defined. These sections are chosen keeping in mind the areas where user feedback is required [4]. The common sections applicable to all functioning websites include:

• Site Layout Features:

These are the features that help in layout management issues of the website [2]. Review of these features is needed as it helps in identifying issues in browsing the website, organization of the different elements of the website and the inconvenient features that the user has to encounter while using the website. The review of the following features is necessary:

- Site Map: It gives the user an easy outline of the site and its features.
- Search Feature: It is an integral part of each website for browsing and locating services.
- Arrangement of Sections: A website should be well-organized with respect to its sections and every section should be conveniently viewable and accessible.
- **Ease of Use:** The complexity of surfing the website as well as browsing, performing operations and transactions.
- User based services:

These services provide more intrinsic feedback based on each user experience individually. The feedback of these services is essential to make the website more user-centric. The services included for the review are:

- **Social Integration:** This section will focus on integration of different social media accounts of the users. This is achieved by introducing Share buttons, twitter feeds etc.
- **Relevance of Recommendations:** This section focuses on how relevant the recommendations showing up on the user's page are.
- **Email Newsletters:** Newsletters are a way to keep users from forgetting about the website by keeping them in the loop about new updates.
- **Helpful ads:** Since ads are an inevitability, it is important to have meaningful and helpful ads relatable by the users.
- Interactive forms & Chat boxes: Chat boxes and online forms provide a way for the user to ask question and make inquiries. The more conveniently a user can get in touch with customer care, the more likely he is to visit the site again.
- Quality & Support:

These features help in management of quality, maintenance and support of the products and services provided by the website. Review of these features ensures that users are satisfied with the services provided by the website

- **Quality:** Quality plays an important role in gaining a user's trust and time. The quality of the content being shown on the site and the products that were delivered has to be top notch.
- **Maintenance:** Maintenance of the website during high network activity is critical. It should be able to handle slow internet connections, high traffic, low bandwidth and cyber-attacks.
- **Support:** Customer support remains a priority. Exchange, Returns, Refunds should be easy for the user.
- Secure Services and Privacy: Services like online transactions and storing private information of the users has to have a secure portal and should be resistant to all kinds of attacks.

• Comments:

Comment section is slowly becoming an irreplaceable part of decision-making process for a user. It is important to customize it in the best way possible to provide more inputs about the products, services or functionalities that are not covered in the feedback. It provides users the freedom of expressing their views about any area not restricted to the ones covered in the feedback forms.

For example: In an e-commerce website, the basic sections for which the user feedback questions will be generated are as follows:

1. Site layout: It covers features of site like searching a product, comparing various products, browsing categories, placement of sections like trending products, offers etc.

E.g. Do you face any inconvenience in browsing different categories?

2. User based services: These include the feedback for the website's social networking accounts, the newsletters of latest offers, how relevant the recommendations of products are and how interesting the ads are.

E.g. How relevant are the offers recommended to you?

3. Quality and support: These include factors like, the quality of the products being sold, the maintenance of the website during flash sales, the support provided by the website for refund, exchange and return, also how the complains about services are handled and is the privacy of the user being compromised in any way.

E.g. In how many days is your complaint addressed?

4. Comments: This covers additional comments that the user can utilize to bring an issue to notice.

3.2. User Identification and Classification

The more experienced the user is with the site, the better feedback he can provide. In the proposed system a module is provided for classification of users based on their usage of the site.

The classification is based on the Pareto Principle [5] as, by obtaining and analyzing the feedback of top 20% users of the website, 80% of the related user patterns of feedbacks will be covered.

To identify these top 20% users and provide more weightage to their reviews, all users are classified into discrete levels from 1 to 5 as classifying users into 5 levels will make it easier to find the top 20% users. As the number of users is large and their usage varies, at least 5 different levels are needed to classify them. In this module, the user is identified and classified into one of the predefined levels. The predefined levels are based on the following factors:

- 1. How active the user is on the website.
- 2. What features/services he has used.

3.2.1. To determine how active the user is:

How active the user has been is calculated based on three factors: number of days since the user created his account, number of times he logged in since then and average time user spends browsing the website. Let's say A is the activeness of the user, then:

A = log(n * l * t)

Where, n = number of days since the user joined

l = number of times user has logged in to the website

t = average time user spends browsing the website

3.2.2.To determine features/ services he has used: Points have been assigned to each feature that the user has used such that if he has higher points in the services used or is more active on the website, his level will be higher.

How many features and services user has used is calculated based on these points. These points can be assigned by the website developers according to their needs. For example, in an e-commerce website points to the following features could be assigned according to Table 1. Let's say S quantifies the services used by user, then

Table 1	
Features	Points
Basic functioning features:	
Browsing the products	15
Searching for a product using search bar	15
Using filters while viewing products (price, size etc.)	25
Comparing different products	25
User Registration features:	
Creating account	10
Storing user details	15
Transaction features:	
Adding a product to wish list	25
Adding a product to cart	35
Placing an order	70
Completing a transaction	90
Total number of products purchased (p)	10*p
Social Media features:	
Sharing products on different social platforms	60
Referring a friend	60
Customer Review features:	
Providing reviews	40
Providing image/video reviews	60
Asking a question	20
Answering a question	50
Rating a product	30
Customer Care features:	

Exchanging a product	20
Returning a product	10
Complaints	10

S = total points of services used by user/100

Here S is calculated by adding the points according to the features user has used divided by 100 to get the value of Y in the range of 1 to 10.

Now, the score of the user is calculated based on these two quantities.

Hence,

$$Score = (A + S)/4$$

SCORE	LEVEL
0-0.99	1
1-1.99	2
2-2.99	3
3-3.99	4
>4	5

Table 2 shows the score and the corresponding levels.

Based on the above algorithm, the users are classified into levels ranging from 1 to 5 depicting their interaction with the website, with level 1 corresponding to an elementary user making use basic use services and level 5 corresponding to an experienced user making extensive use of services.

This classification is then used to generate questions of each section and assign the complexity to them. Greater the level of the user, the more detailed and advanced is the question. Higher level feedbacks are given more weightage since they are more reliable as compared to lower level feedbacks.

This works on the principle that a more experienced user who is well versed with the functionalities of the website can provide better and detailed insights. Higher level of classification denotes better understanding of the services and hence will be used to extract deeper and more accurate insights to identify problems and solutions in each section. For every level, questions in each section are different to extract more information out of a diverse group of users. Higher level questions are more elaborate whereas lower level questions are more basic in nature. The system adapts itself according to the identified user, thus making the feedback adaptive.

Updating the Levels: With time, user levels are updated as the classification process will be repeated regularly. Thus, with more comprehensive usage, the level of the user will increase. Also, for a website, after a specific time, the parameter values for classification level criteria can be upgraded to keep the classification up-to-date.

3.3.Determining Usability of Feedback

When the number of users is large, not every feedback can help in providing valuable insights, thus a filter needs to be added to discard the feedbacks that do not provide new inputs.

For this, all answers are compared to the predefined default values for every question, the matched ones are discarded as they provide no new insight and the remaining answers are sent for analysis. This is done because the set of default values is selected keeping in mind prerequisite knowledge and commonly existing patterns, thus only that data is required which provides new insights and undiscovered patterns.

Figure 2 shows the process of determining usability of the feedback form.

For example: In a feedback of 15 questions, if the user has kept the default values intact for 10 questions, then only the 5 answers that provide some new knowledge for analysis will be considered and rest will be discarded.





Weightage of the feedbacks

More weightage is given to the user feedbacks of higher levels, i.e. level 5 feedbacks will be given more weightage than level 4 and so on. This is based on the fact that level 5 users have better understanding of the website functionalities than lower level users.

By assigning increasing weightage corresponding to the increase in levels, identifying issues in feedback will become easier. Feedbacks from level 1 provides a general trend and feedback from successor levels provides a clearer view as the levels increase, with level 5 feedbacks pinpointing to the problem or the ideal solution.

Following analysis of user feedback

Once all the feedbacks have been analyzed and combined with user monitored behavior to generate patterns through sentiment analysis [6], if for any of the basic sections the following conditions exist:

- 1.No new patterns can be identified
- 2. Problem cannot be correctly identified
- 3. Solution cannot be rightly pointed out

More efficient information needs to be extracted out of the user feedback and thus the questions in the section need to be,

A. Reframed: The same question is framed in a different manner or form of the question is changed. The form of the question corresponds

to selection from predefined options, manual comments, single option selection, multiple selection option, etc.

- B. Elaborated: The same question is elaborated in the form of sub questions or specific area to which the question is pointed is mentioned clearly.
- C. Replaced: The existing question is replaced with a clearer, detailed and better suited question.

For example: If the social integration of the e-commerce website is weak and the original questions in the feedback is:

How useful is the information about the latest offers on our Facebook and Twitter accounts?

If no specific issue can be identified from this question, the website needs to adopt one of the following measures:

i. Reframe:

Do you think our Facebook and Twitter accounts are up to date with the latest offers?

- ii. Elaborate:
 - Do you find our Facebook and Twitter accounts useful in providing details of the latest offers?
 - Do you follow the links shared on our Twitter account about latest offers?
 - Are the queries and comments about the latest offers addressed immediately on our Facebook page?
 - How informative are the advertisements uploaded on our Facebook page?

iii. Replace:

Which of these social mediums helps you the most in providing information about the latest offers and upcoming sales?

4. RESULTS

The e-commerce example has been considered to expand on the idea of adaptive feedback below:

• New User

Let us take an example of a newly joined user X. Let,

n, number of days since the user joined = 7

I, number of times user has logged in to the website = 10

t, average time user spends browsing the website = 15 min.

Since $A = \log (n * l * t)$

Thus $A = \log (7 * 10 * 15) = \log (1050)$

A= 3.0211

Say he has browsed the website, searched for specific products, applied filters while viewing products, created an account and added a product to wish list. According to Table 1, his total points for all the features used will be

$$S = (15 + 15 + 25 + 10 + 25)/100$$

S = 0.9

Now to calculate which level he belongs to, the formula (number) is used.

Score = (A + S)/4 = (3.0211 + 0.9)/4

Thus Score = 0.98

According to Table 2, this user X belongs to level 1 and thus that set of survey questions will be provided to him.

• Old Active User

Now let us look at an active user Y who has been using the website for a long time. Let,

n, number of days since user joined = 700

I, number of times user has logged in to the website = 624**t**, average time user spends browsing the website = 45 min Since A = log (n * 1 * t) Thus A = log (700 * 624 * 45) = log (19656000)

Thus $A = \log (700 * 624 * 45) = \log (19656000)$ A = 7.294

Say he has used all the features mentioned in Table 2 and total number of products ordered by him is 25. So his total points for all features used will be

S = (935)/100 = 9.35Now his score is calculated: Score = (7.294 + 9.35)/4 = 4.16 According to Table 2, this user Y belongs to level 5.

Figure 3 and 4 show how the feedback form will adapt itself to the user classified level. Since user Y is more experienced about using the website than user X, Y's feedback is more comprehensive and will be given more weightage as compared to that of X.

12	3	4 5	5							□ Yes					
0 0	0	0 0	_] No					
How would you rate the quality of the products sold by our								ur website?	vehsite? * Do you find our Facebook page informative? *					ve? *	
1 0 0 1 5 (7 0 0 10															
1 2	3	4 5	0	1	/ 8 9 10										
0 0	0 0 0 0 0 0 0 0 0														
How wo	How would you rate the customer care provided by our we						ebsite? * D	ite? * Do you face the "Server has Crashed" problem often? *							
Good Good	i								Rarely						
🗆 Bad									0 [ften					
Very	Good								N	lever					
Very	Bad								Always						
Are the	prod	ucts red	comr	nend	led to	you rele	vant to your n	eeds? *							
□ Yes															
🗆 No															
□ Ofte	n														
Rare	ely														
						Fig	gure 3(Lev	el 1 ques	tio	ns)					
Which of the fo	lowing	features	do ye	ou fin	d difficu	ult/ inconv	enient to use? *	How often	are	Server ha	s crashed	' proble	m faced by	VOII: *	
Searching a product								During.	Col	05				,	
Comparing a product Received Categories															
Finding relevant offers															
Applying filte	ers														
Which of the following criteria would you prefer for the products recommended to you?*							How would	a you describe our Facebook page for the followin Yes No			e following ca No	ategories? *			
Past purchases							Up to Da	te		0		0	0		
Past browsing history						Informat	Informative			0		۲			
□ Similar to the viewed items						Interactiv	Interactive o			0	0				
□ All of the above						Relevand	Relevance of Content O				0				
								Spam Co	onte	nt	0		0		
How would you classify the ads on our website? *					How would you r	ate th	e quality of the pro	oduct according to	these criteri	as: *					
	1		2		3	4	5			1	2	3	4	5	
Relevant	0		0		0	0	0	Similarity with online descript	tion	0	0	0	0	0	
Offensive	0		0		0	۲	0	the product	of ith	0	0	0	0	0	
Spam	0		0		0	0	0	the product	IUI	0	0	0	0	0	
How would	d you	rate t	he c	ust	omer	care of	f our website	e for the fo	ollo	owing: *					
1 2						3 4			5	5					
24*7 su	ppor	t			0		0		0 0			0	0		
Friendlin staff	ness	of			0		0		0 0			0	0		
Response time for complaints					0		0		0		0		0	0	

How convenient and accessible are the features of this website? * Do you find the advertisements on our website helpful?*

Figure 4(Level 5 questions)

5. CONCLUSION

User feedback is one of the most crucial feature for improvement and development of a website, it works as a first-hand experience of the people for whom the website is built. Thus, they can provide the best insights for identifying problems as well as solutions. In this proposed work, Pareto Principle has been used to classify users on the basis of their experience and usage of the site. This classification produces an adaptive system where the complexity of questions and the weightage provided to the feedback is based on user interaction with the site. The concept of giving more weightage to the feedback of the most interactive and experienced users proves to be a benefit to both the users and the website, thus making the system more reliable than the existing one. This system not only proves to be an efficient method for feedback collection but also provides a quick filtering method for the feedback to make the analysis and extraction part easier.

6. FUTURE SCOPE

With big data being present all around us, an adaptive user feedback system which extracts only necessary and filtered data is the need of the hour. This feedback system can be implemented not only in E-commerce sites for improving user shopping experience but also in Social networking sites for engaging the user and identifying the unpopular features and in organizational websites for supervising the functioning of the organization. Additionally, a third party software can be created based on the proposed system that provides adaptive feedback generation to websites. Thus, this system can conveniently be implemented in the existing websites to make an optimal use of the user feedbacks.

7. REFERENCES

- Gabriella Pasi, "Implicit feedback through user-system interactions for defining user models in personalized search", 6th International Conference on Intelligent Human Computer Interaction, IHCI 2014
- [2] Yi Zhu Li He, Xiaojun Wang, "User Interest Modeling and Self-Adaptive Update Using Relevance Feedback Technology", International Workshop on Information and Electronics Engineering, 2012
- [3] Farid Aulia Tanjung, Wawan Dhewanto, "Formulation of E-Commerce Website Development Plan Using Multidimensional Approach for Web Evaluation", The 5th Indonesia International Conference on Innovation, Entrepreneurship, and Small Business (IICIES), 2013

- [4] Gerson Tontini, "Identifying opportunities for improvement in online shopping sites", Journal of Retailing and Consumer Services, Volume 31, July 2016
- [5] Qing Li, "E-commerce of Eco Bags on Basis of Pareto Improvement", Chapter "Innovation in the High-Tech Economy", Part of the series "Contributions to Economics", November 2013
- [6] Ronen Feldman, "Techniques and applications for sentiment analysis", Communications of the ACM, Volume 56 Issue 4, April 2013
- [7] T Prasad, V Dilip Kumar, "A New Approach for Information Collection from Web With Ontology", International Journal of Applied Engineering Research, Volume 11 Number 9, 2016
- [8] Madhuka Udantha, Surangika Ranathunga, "An Episodebased Approach to Identify Website User Access Patterns", International Conference on Pattern Recognition Applications and Methods, January 2016
- [9] Dora Dzvonyar, Stephan Krusche, "Context-Aware User Feedback in Continuous Software Evolution", Proceedings of the 1st International Workshop on Continuous Software Evolution and Delivery. IEEE/ACM, 2016