

Measuring Users' Contribution in Wikis

Timothy Yau, Priyanka Pareek,
Ofer Arazy, Eleni Stroulia

**Computing Science and School of Business
University of Alberta**

Broad Hypothesis

According to the Collective Effort Model (CEM) by social psychologists Karau and Williams,

motivation to contribute increases when

- (a) Individuals see the value and uniqueness of their own efforts with respect to the group outcome, and*
- (b) Individuals know that others can see their contribution*



Making visible a user's impact to the wiki will increase users' motivations and participation levels

Research Questions

1. How to define a “user’s contribution” (e.g. based on his effort? Contribution’s impact on others?)
2. How do we measure user’s contribution?
3. How do we make the relative contributions of user visible, such that:
 1. The representation is concise and easy to understand
 2. The measurement algorithm is hidden (making it difficult to manipulate the measure).
4. How is behavior affected when contribution metrics are presented on corporate and public wikis?

Previous Related Work

- Vassileva:
 - simple measures of contribution, such as the number of contributed items (i.e. edits); captures quantity (i.e. user's effort) but not quality
- Rashid:
 - In the MovieLens online movie-recommender community, users were more likely to rate movies when the system indicated a high “anticipated benefit” for doing so, especially if the beneficiaries were users with similar movie interests
- Arazy
 - For corporate wikis, users expect that adding an attribution feature would encourages them to contribute more

Research methodology

1. Design

- Defines metrics of user contribution
- Design algorithms to calculate each metric

2. Test

- Select a sample of wiki pages
- Run algorithms to extract metrics of user contribution
- Manually analyze the sample wiki pages, and rate the perceived relative contribution of each user
 - Categorize contributions to a number of classes (adding content, synthesis of existing information, adding links to other pages)
- Compare the automatically calculated metrics against the manual evaluations, and produce a formula that takes the various metrics as an input and generates an overall measure of users' contribution

3. Design user interface to present users' contributions

Terms and Definitions

- ***Wiki page***
 - the document corresponding to a specific wiki page
- ***Version***
 - the result of an edit action on a wiki page
- ***Release***
 - the last of a sequence of same-person versions
- ***Contributors (release)***
 - the set of wiki users who own some element of a page release
- ***Editors (page)***
 - the set of wiki users who have edited the page at some point in its history

Intuitions

- A user **contributes to a page** if:
 - He spends significant **effort**
 - Edits/comments the page frequently
 - She “**owns**” a portion of the current page release (in terms of sentences she has originated)
 - His contributions have **impact** on others
 - The contribution attracts readers
 - The contributions sparks a discussion
- A user **contributes to the wiki** at large is:
 - He contributes significantly to pages that are **important**
 - A page is *important* if it is a hub or it is frequently visited

Measuring User's **Contribution to a Page**

- *Sentence Ownership*: Percentage of sentences/fragments created by the user (see Sentence-level User Contribution) that remain at current release
- *Number of edits*: Edits made by user
- *Number of comments*: For each user, the number of comment (or discussion page) additions over the history of the article
- *Number of links created by the user*: including both hat internal and external link
- *Classification of change type*: automatically analyzing user's contributions and classifying them as: content change, structure change, or both (Jacqueline is doing this)

Measuring Page Importance

- *Frequency of page visits by users:* MediaWiki maintains a hit count for each page, updated periodically.
 - Our algorithm calculates the number of hits over last 30 days
- *Incoming internal links:* Pages within the wiki linking to this article.
 - This feature is provided in MediaWiki
- *Incoming external links:* the page's ranking according to Google (if the wiki is in the public domain).
 - Not implemented yet, however, there are some sites such as SEOpen.com that are able to retrieve pageranks from Google, but according to [this site](#), the technique may violate Google's terms of service
- *Rating of content:*
 - we plan to implement a functionality to allow users to rate

User's Contribution to the Wiki At Large

- Based on User-Page relationship
 - General
 - Using the most recent version of each page
 - Weighting each page's contribution by the importance of the page
 - Measurements
 - Content: sum of user-owned fragments on a page
 - Discussion: sum of user comments on a page
 - Diversity: number of different pages edited
- Based on user-user relations: user's role in the social network;
 - Definitions
 - Two authors are connected if they have edited the same wiki page
 - An author is connected to a reader who has read a page a contributed to
 - Metrics
 - Degree centrality: the extent to which the authors is linked to other authors and readers
 - Betweenness centrality: the extent to which the author bridges disconnected groups of authors

Manual Evaluation of Articles

- 20 wiki pages
- A page-level evaluation of user contribution
- A single evaluator using pre-defined guidelines
- Evaluator's perceptions based on:
 - Overall perception of top contributors to a page
 - Analysis of the page history log, examining each edit
 - The type of the edit
 - *Structural Change*: re-organizing text or creating new sections
 - *Correction*: word-level correction
 - *Add*: Adding text to article at word level
 - *Delete*: revert or undo operations
 - *Internal References*: adding internal links
 - *External References*: adding external links

Example: Sentence Ownership

Each colour represents a different user

[edit]

Defense

Aikido techniques are largely designed towards keeping the attacker off balance and locking joints. Manipulation of *uke's* balance by entering is often referred to as "taking the centre". It is sometimes said that Aikido contains only defense, and the attacks that are performed are not really attacks. Habj 14:40, 28 December 2004 From a historical perspective this claim is questionable, but many if not most aikido practitioners focus on defense techniques as the focus of their training. Much of aikido's repertoire of defenses can be performed either as throwing techniques (*nage-waza*) or as pins (*katame-waza*), depending on the situation.

Each technique can be executed in many different ways. For example, a technique carried out in the *irimi* style consists of movements inward, toward the *uke*, while those carried out in the *tenkan* style use outward sweeping motions, and *tenshin* styles involve a slight retreat from or orbit around the point of attack. An *uchi* style technique takes place in front of *uke*, whereas a *soto* style technique takes place behind him; an *omote* version of a technique is applied directly, whereas an *ura* version is applied using a turning motion; and most techniques can be performed when either *uke* or *nage* (or both) are kneeling. Thus from less than 20 basic techniques, there are literally thousands of possible actions depending on the attack and the situation.

There is also the matter of *atemi*, or strikes employed during an aikido technique. The role and importance of *atemi* is a matter of some debate in aikido. Some view *atemi* as strikes to "vital points" that can be delivered during the course of a technique's application, to increase effectiveness. Others consider *atemi* to be methods of distraction, particularly when aimed at the face. For instance, if a movement would expose the aikido practitioner to a counter-blow, he or she may deliver a quick strike to distract the attacker or occupy the threatening limb. (Such a strike will also usually break the target's concentration, making them easier to throw than if they are able to focus on resisting.) *Atemi* can be interpreted as not only punches or kicks but also, for instance, striking with a shoulder or a large part of the arm. Some throws

User Contribution to the Wiki at Large

All Users

No groups selected; displaying all users.

Username	Number of Articles Edited	Number of Edits	Sentences Owned	Comment Sentences	Internal Links
	167	801	2329		315
	176	511	2144	2	307
	44	368	1144		35
	78	362	1328	6	106
	39	274	376	601	5
	20	184	417		13
	19	161	400	20	14
	18	156	431		10
	43	149	598		36
	16	145	390		10
	19	131	718	14	14
	1	112	19	113	
	13	78	272		12
	45	70	204		47
	2	62	20	160	

User Contribution to the Wiki at Large

Details for one specific user

Main Namespace Contributions

Page Title	Sentences Owned	Percentage of Page	Edits	Links Owned	Comment Sentences
	33	100.00	92	1	
le	72	100.00	50	1	541
	136	84.47	18	1	
	1	1.39	12	1	
	6	75.00	5		
	6	46.15	5		
	5	83.33	5		
3	3	75.00	5		5
	2	100.00	4		
)	1	100.00	4		1
)	1	100.00	4		2
P	1	100.00	4		4
	1	100.00	4		1
	1	100.00	4		2

Page Importance Metrics

(& additional page statistics)

Page	Edits	Views	Edits (30 day)	Views (30 day)	Incoming Links	Sentences	Outgoing Links	Outgoing Links (broken)	Comment Sentences	Contributors
🔗	1004	6	0	0	0	230	0	0	608	386
	984	10	0	2	0	274	0	0	724	450
	893	17	0	0	0	568	0	0	475	530
	633	9	0	0	0	150	0	0	249	327
🔗	449	5	0	0	0	312	0	0	257	290
	374	10	0	0	0	306	0	0	133	206
	362	2	0	0	0	137	0	0	135	203
al_equation 🔗	321	2	0	0	0	341	0	0	301	151
s 🔗	316	8	0	0	0	239	0	0	162	177
	255	1252	0	0	152	3	3	0	6	10
	202	6	0	0	0	130	0	0	101	108
e_Control_and_Prevention 🔗	175	6	0	0	0	48	0	0	33	86
	126	1991	0	0	69	149	62	3	0	1
	114	5	0	0	0	41	0	0	37	84
	110	5	0	0	0	49	0	0	104	81
	93	38	0	0	1	33	1	3	0	2
	88	3	0	0	0	36	0	0	13	65
🔗	70	4	0	0	0	121	0	0	42	49
	67	1892	0	0	1	40	0	0	0	3
_Project 🔗	61	1500	0	0	1	108	0	0	0	2

Implementation Status

- Algorithms and scripts to compute metrics have been implemented
- Plug-in to MediaWiki:
 - Computes metrics
 - Provides rudimentary presentation of metrics (font-highlighting + tables)
- On-going research:
 - Presentation: refinements
 - Derivation of a formula for calculating the overall measure of users' contribution (taking the various metrics as an input)
 - Based on comparison of automatic algorithm and human perceptions

Conclusion

- In-depth analysis of user contribution to wiki, considering various user activities
- The impact (as far as affect on users' motivations and participation) of displaying user contribution measure depends on the:
 - Accuracy of the algorithm
 - Graphic representation
- A single measure of user contribution is advantageous
 - Easier to understand
 - Difficult to manipulate
- This method should work well for corporate wikis, where there are benefits for demonstrating performance