Analyzing How Users Utilize 'Riff' for Collaborative Searching and Sharing Contents for Social Learning in So.cl

Erik Choi erikchoi@gmail.com
Vanessa Kitzie vkitzie@gmail.com
Chathra Hendahewa chathra.hendahewa@gmail.com

Rutgers, The State University of New Jersey

Introduction

Microsoft So.cl combines the collaborative features of a social network and the capabilities of a search engine to create a unique social search environment. The activity of riffing provides a unique twist on a typical search results page since users can actually contribute to the results of another persons' search string, and/or annotate their own searches, also providing multimodal content such as image, video, and text within one results display; as well as user generated metadata via comments and tags. These collaborative and search features give so.cl the potential to promote discovery through interaction with search result content. Unfortunately not many so.cl users exhibit riffing behavior. For this reason, we explored behaviors of users who had riffed in order to further examine their behaviors, focusing also on some of the top riff posters, to determine both what behaviors and post activity might be popular among these users. Results from this analysis indicate that overwhelmingly users who riff on others' posts also tend to create their own riff postings, as well as the fact that as users tend to become more acclimated to the so.cl service, they also tend to riff more. These findings can be used to determine implications for pushing certain behaviors within so.cl to promote riffing behavior.

Riffing

Of the several features that Microsoft So.cl provides for sharing and organizing content, our specific focus is on "riffing" behavior. Riffing behavior differentiates Microsoft So.cl from other social network sites such as Facebook, Twitter, or Pinterest. When you "riff" on a user's post, you have the ability to re-post a web object or piece of information that someone has shared and add your own annotations including comments and tags. People who look at your riff have the ability to see all of the riffs emanating from the original post thereby exposing them to research that has already been conducted on their topic of interest.

Incorporating riffed content into a results display not only allows people to be serendipitously exposed to user curated content, as well as pages ranked via a conventional web algorithm, but also allows users to search different content (e.g. text, images, gifs, video, audio) within the same results page. This differs from other commercial search engines in which the user must determine the type of content that they would like to search (e.g. video, images, text). By predicting the content and type of information that users are likely to riff, we can make some inferences about how well two users with similar topical interests can benefit from collaborative learning such as following each other, sharing content and by riffing each other's posts.

Research Questions

The study focused on determining how riffing relates to other behaviors, specifically providing feedback on riffed posts and posting other online content. Specifically, we would like to explore the following research questions:

Choi, E., Kitzie, V., & Hendahewa, C. (2013). Analyzing how users utilize 'riff' for collaborative searching and sharing contents for social learning in So.cl. *iConference 2013 Proceedings* (pp. 1114-1117). doi:10.9776/13531
Copyright is held by the authors.

RQ1: What are the textual and non-textual features that contribute to a specific piece of content getting riffed?
RQ2: Who are people who frequently riff? How are they the same or different from other users who may not riff as much?
RQ3: How does so.cl activity relate to riffing behavior over time?

Data Analysis

The data was extracted from so.cl Behavior Data files, which were provided by Microsoft, with a total number of 313,745 unique users. Out of this user base, we extracted 4,042 unique user ids of people who have exhibited riffing behavior, or 1.29% of the available data. From here, the study focused on the riffed posts a user created, other riffed posts the user might have provided some form of feedback (e.g. link, comment) on, as well as any other content that the user posted within the so.cl environment.

From a random sample of N=1,000 UserIDs, the study found that the number of other online posts outside of riffing (n=582,598) greatly outnumbered riffing content (riffed posts, n=12,079; feedback on riffed content=11,405) among users who had displayed riffing behavior. The most popular type of online behavior is posting a photo (n=246,684), which accounts for 40.7% of the total content measured - riffed posts, feedback on riffed content, and other posts (See Table 1).

Table 1. Descriptive Statistics on behaviors of users who riff

	Total Riffs	Own Riffs	Riffs to Others	Total Feedback Riffs	Own Riff Feedback	Riff Feedback to Others	Total Other Online Posts
N = valid 1,000	1000	1000	1000	1000	1000	1000	1000
Mean	12.08	10.21	1.87	11.41	1.69	9.71	582.60
Median	1.00	1.00	0.00	1.00	0.00	1.00	31.00
Mode	1	1	0	1	0	1	3
Std. Deviation	58.854	49.828	10.967	55.123	9.090	48.226	2508.590
Variance	3463.75	2482.78	120.27	3038.52	82.63	2325.78	6293025.38
Skewness	8.67	8.60	9.20	8.87	9.81	9.14	10.04
Std. Error of Skewness	.077	.077	.077	.077	.077	.077	.077
Kurtosis	88.99	85.56	96.75	94.49	114.37	100.90	137.46
Std. Error of Kurtosis	.155	.155	.155	.155	.155	.155	.155
Minimum	1	0	0	0	0	0	0
Maximum	793	653	163	764	145	681	45963
Sum	12079	10207	1872	11405	1691	9714	582598

A linear regression was performed using number of riffed posts as the dependent variable in order to better examine what types of content (feedback on own riffed content, feedback on others' riffed content, other post types - photo, note, link, comment, tag, video, news, document) were making the most significant contribution to explaining the variability of total number of riffed posts. Surprisingly, given its low sample size (n=9,714) when compared to other online content, feedback on another user's riffed content had eight or more times the beta value size of any other factor included in the regression (β=0.861). A subsequent stepwise regression revealed that feedback on another user's riffed content explained 98.2% of the variability of number of riffed posts (r=0.991). On the other hand, photo posts constituted one of the only variables that did not make a significant contribution to the model (p=0.477). This is particularly surprising given that the large sample size of photo posts alone would suggest statistical significance (See Table 2).

Table 2. Result of linear regression on number of riffed content

		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-.055	.173		-.321	.748
	Feedback on Own Riffed Content	.864	.036	.133	24.191	.000
	Feedback on Other Riffed Content	1.051	.007	.861	143.647	0.000
	OtherPostPostType_Note	0.20	.004	.046	4.836	.000
	OtherPostPostType_Photo	.000	.000	-.004	-.712	.477
	OtherPostPostType_Link	-.011	.003	-.057	-3.746	.000
	OtherPostPostType_Comment	-.051	.006	-.154	-8.760	.000
	OtherPostPostType_Tag	0.21	.002	.068	10.139	.000
	OP_Like	0.10	.002	.075	4.490	.000
	OP_Video	-0.33	.005	-.095	-7.225	.000
	OP_VideoParty	-.074	.047	-.005	-1.575	.116
	OP_News	.072	.009	.075	8.386	.000
	OP_Like-Comment	.055	.004	.093	14.071	.000
	OP-Doc	-.750	.346	-.006	-2.169	.030

The variances for the total number of riffed posts, feedback on riffed content, and other online posts are similar relative to the sample size and positively skewed, which indicates that a smaller number of users provide the majority of the content. For this reason, the study decided to further examine top users and their riffing behaviors. The top 25 users across the following categories were first focused: total number of riffed posts, total number of riffed feedback, and total number of other online posts. The study found a 92% agreement between the top 25 users among riffed posts and feedback on riffed posts, and ~60% agreement between top users of these categories and top online posters. Across each category, users also posted ~60% of the total content. From here, top 25 posters of riffed content were focused to compare their cumulative amount of so.cl activity to the other 9,975 sampled users. Independent t-tests assuming non-homogeneous variances indicated that there are a statistically significant differences between the amount of so.cl activity among the top 25 posters and the other users ($p<0.01$) for all areas except for amount of postings of documents (this could be explained by the relatively small sample size, n=17).

Finally, the study attempted to examine how the activity of our top users has changed over time. Did users tend to create riffed posts more, less, or the same amount after they became acclimated to the so.cl site? To measure this, the study first calculated the median time and date of online postings for each user; the median was chosen as an indicator of acclimatization. Then, the study also counted the amount of riffed posts they created before and after this time period. Among our top 25 users, findings indicate that there was a 66% increase of riffed posts (n_{before}=2,981; n_{after}=4,954) created post-acclimatization to interacting within the so.cl environment, as indicated by online posts. This suggests that among top users, the more time they spend on the site, the more they create riffed posts. Further study could be performed to determine when rate of riffed posts created plateaus and how this might correlate to other post activity within so.cl.

Discussion

One key finding here is the link between a user creating a riffed post and providing some sort of feedback (e.g. comment, link) on another users' post. This suggests that promoting both activities will marry the objectives of so.cl to both push the social interaction aspect via riffing on others' posts, as well as the customization aspect of interacting with a search results page via riffing. Future studies should be performed in order to further explicate this link and perhaps measure the effect of interventions meant to enhance these behaviors.

Another finding examined the types of users who frequently riffed. We see that their so.cl behaviors as a whole, both riffing and creating other online posts, are significantly different from the rest of the so.cl users who have riffed. Over time, the study also noticed that the frequency of these behaviors tends to increase. These findings are packed with ideas for future study, including a longitudinal analysis of riffing use over time, as well as how behaviors of top users might differ from other users.

A limitation of this study was that it did not focus on people who did not create riffed content. Future studies could use information about these people in order to create a logistic regression model that could indicate what behaviors, user characteristics, and textual/non-textual features of content shared, contribute to a person's likelihood to riff.

Conclusion

There is a hope that findings will inform the creation of a push system that attempts to diffuse information across networks by suggesting specific content to be riffed to other users. This kind of information suggestion is important, since it can lead to discovery of new information that is more likely to be of interest to a related user that has common interests with the person who riffed the content, but also brings a certain amount of uniqueness to the content based on the looser social relation to someone who has the potential to be outside of the user's core network.

Riffing provides a unique way of which the users are able to interact with each other based on their specific interests within Microsoft So.cl network. At this instance, investigating user behaviors related to riffing constitutes the first step for gaining a better understanding of how So.cl users develop social relationships and collaborate. We hope that findings will inform what kind of content is riffed and encourage users to collaboratively search and share content by developing rich user interactions which would pave the way in promoting a mutually beneficial social learning paradigm.