Dynamics of Giving up in Network Search
Group #23

Yonathan Perez
yperez@stanford.edu

Nicholas Shelly
nshelly@stanford.edu

December 10, 2012

Abstract

In this project, we look at why humans give up in network search, and more specifically in Wikipedia navigation. We aim to find features of the mission, the path traversed and their underlying graph structure, which distinguish games that are not completed.

In comparing successful and unsuccessful searches, the paper first examines a priori mission characteristics, that is properties that that depend only on the mission (a pair of start and goal articles), including the shortest path length, the centrality of the goal node, and lucrative ratios. Second, we look at properties that depend on the user's selection of nodes along the path. We examine measurements that depend on both on the mission and on the user and the path the user decides to pursue. Such measurements include the evolution of article properties along the search path, and properties of the subgraph defined by the search path.

We show that even before starting out on the search, certain missions are already more likely to lead to failure, due to a greater shortest path length from start to target nodes, lower lucrative ratio, and a lower PageRank for unsuccessful targets. Along the path, the betweenness centrality of the path's edges also contribute toward the success or failure. Straying to far from the center, can make it difficult to navigate. Using these parameters, it makes it easier to predict when the user will fail.

1 Introduction

The online game "Wikispeedia" provides insight into how humans navigate an information network by challenging the user to travel from one article to a seemingly unrelated, goal article in as few clicks as possible.[1] The creators of the game focused on how the users successfully progress through this small world network and gained insights in human navigation.

In this project, we look at the proposition from the other perspective: why do humans give up in network search? Understanding the conditions which lead humans to abandon a search, can be useful for deciding when to intervene to aid the user, or in setting realistic expectations that a user can find something. We try to find factors that make giving up on the search more likely: whether it is being stuck in structurally suboptimal part of the network, or simply tackling an inherently difficult mission with a goal far from the start point.

Similar to the work done in [1] and [2], we gather our information from traces of real Wikispeedia games. Through exploratory measurements, we aim to find which features of the underlying network structure distinguish the finished and unfinished game trajectories.

2 Related Work

Wikipedia encompasses a small-world network that can be easily navigated in just a few clicks, due to its low clustering coefficient and hub nodes with high degree. Earlier work has provided models for efficiently navigating these networks. Wikipedia seems to fulfill the rank-based friendship model, suggested by Liben-Nowell *et. al.* [3] In the rank-based model the probability of a long-range connection between a node u and a destination node v is proportional to the number of nodes closer to u than v. This rank-based friendship model provides a linear fit between the rank and the probability of long-distance friendship. The authors assert that if a person has no friends closer to the destination, then the person simply "gives up" in a message passing task. In this project, we will look beyond the sheer distance factor, and search for further factors that contribute to search failure.

West, Pineau, and Precup [1] introduce the "Wikispeedia" game. The authors' main goal is to compute semantic distance between Wikipedia articles based on game traces. This is based on the hypothesize that human navigation follows semantic associations based on common-sense knowledge. This means users follow intuitive, not shortest, paths. Humans also try to reach a general concept with many outgoing links (a "hub") as early as possible. This is the initial "getting-away" phase. Later, a player would narrow down on his target, using pages that are conceptually closer and closer to the target. This is the "homing-in" phase. The authors leave open the question about what happens in traces of games that were not completed by the player.

West and Leskovec [2] identify several strategies used for human wayfinding by analyzing real traces of Wikispeedia games. The authors present two goals: one is analyzing search paths and gaining insights on human wayfinding, and the other is predicting the goal based on a partial path traveled. This paper presents some key insights on human navigation strategies. Humans tend to have a faster progress rate, and use more predictable steps, at the beginning of the search (far from the target) and towards the end of the search (close to the target). In long searches, humans tend to have an inefficient circling phase, in which they stay within the same distance from the target and use less predictable steps, before finally homing in on the target. The research also shows that humans include factors of both page degree and conceptual similarity in their wayfinding strategies, but at a varying extent along different parts of the search. The research also shows that a multi-category strategy leads to much better performance than a "simple" strategy (traveling across multiple hubs), suggesting a trade-off between simplicity and effectiveness. Finally, the authors offer a mechanism for predicting a goal based on a path prefix. We follow the analysis done by West and Leskovec [2], but aim at those search missions that were not completed. Over half (54%) of all search missions were canceled before finishing, with a significant drop-out rate of approximately 10% at each step in the path. Similarly, Milgram's experiment showed that only 18% of chains were completed to the target individual.[4]. We will look at how the underlying network structure can contribute to search failure (player drop-out), and reason about why do people give up their search.

3 Findings on Giving Up in Search

3.a Methodology

In this section, we measure and analyze some of the properties that characterize games, or search missions, in which the user gave up before completion. The properties we assessed can be divided into two main categories:

The first category is of features of the mission itself. Properties that belong to this category depend only on the mission (a pair of start and goal articles), and not on the user or the path taken. Thus, such properties can be referred to as *a priori mission hardness*.

	Finished	Unfinished
Avg. SPL	2.88	3.26
Avg. # of SPs	2.97	3.29

Table 1: start to goal shortest paths

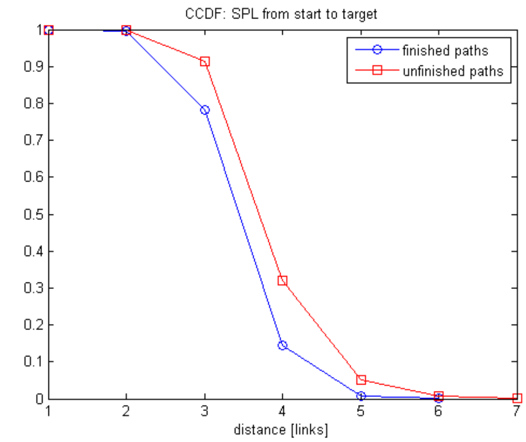


Figure 1: CCDF of SPLs from start to goal article.

The second category is of properties, or measurements, that depend on the entire path taken. Here we discuss measurements that depend on both the mission and on the user and the path the user decides to pursue. Such measurements include the evolution of article properties along the search path, and properties of the subgraph defined by the search path.

3.b Measurements that depend on mission only

Distance from start to target One of the most obvious criteria for *mission hardness* is the distance between its start and goal articles. When comparing the distributions of the shortest path length (SPL) between the start and target articles, we indeed see that these articles tend to be farther away from each other for the unfinished paths. This can be seen in Table 1. We also see that the number of shortest paths from start to goal, as shown in Figure 2, is higher for unfinished paths, which reasons well with the previous finding. The exact figures for the shortest paths from start to goal are listed in Table 1.

Textual distance between the start and goal articles, however, does not seem to have an effect on whether a path is easier to complete. For textual distance between two articles we use one minus the cosine of their TF-IDF vectors (similarity). In almost all missions, the start and goal articles are not semantically related at all – having a textual distance of 1 – whether the path is completed or not.

Centrality of goal Another measure for a mission's

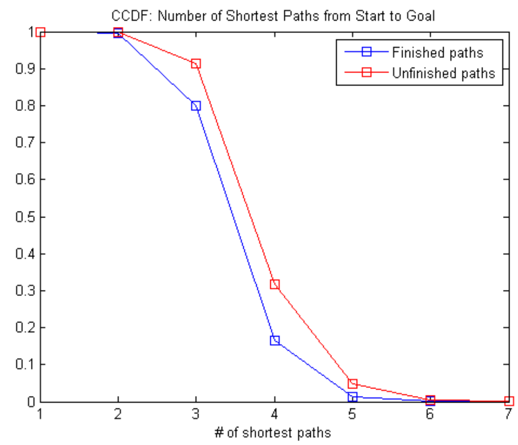


Figure 2: CCDF of # of SPLs.

	Finished	Unfinished
closeness	684	550
PageRank	0.0004	0.00019
betweenness	15234	32354

Table 2: Centrality properties of target article (average)

a priori hardness is the goal article's centrality (or "importance") in the Wikipedia network. There are several ways to define the centrality of a node within a network. We use three notions of centrality:

1. PageRank score of article (with a teleportation probability of 0.15).

2. Closeness centrality of article. A node's closeness centrality captures how close the node is to the rest of the nodes in the network, and can be defined for a target node as:

$$c(u) = \sum_{u \neq v} 2^{-d(v,u)}$$

3. Betweenness centrality of article. We used the definition of betweenness centrality from [5], with the adaptive-sampling based approximation and $c = 5$. The dependency of article $s \in V$ on any other article $v \in V$, where σ_{sv} is the number of shortest paths from s to v, is defined as:

$$\delta_{st}(v*) = \sum_{w:v \in P_s(w)} \frac{\sigma_{sv}}{\sigma_{sw}}(1 + \delta_{s*}(w))$$

The total betweenness centrality of node u is equal to:

$$\Delta_u = \sum_{s \neq v \neq t \in V} \delta_{st}(v)$$

Figure 3 shows that the centrality of the goal node is higher among finished paths than unfinished paths, according to the PageRank and closeness centrality measures. The distribution of goal betweenness centrality is more ambiguous. The average betweenness centrality of the goal article in unfinished games – to which a user never actually reaches – is higher than its equivalent among finished games.

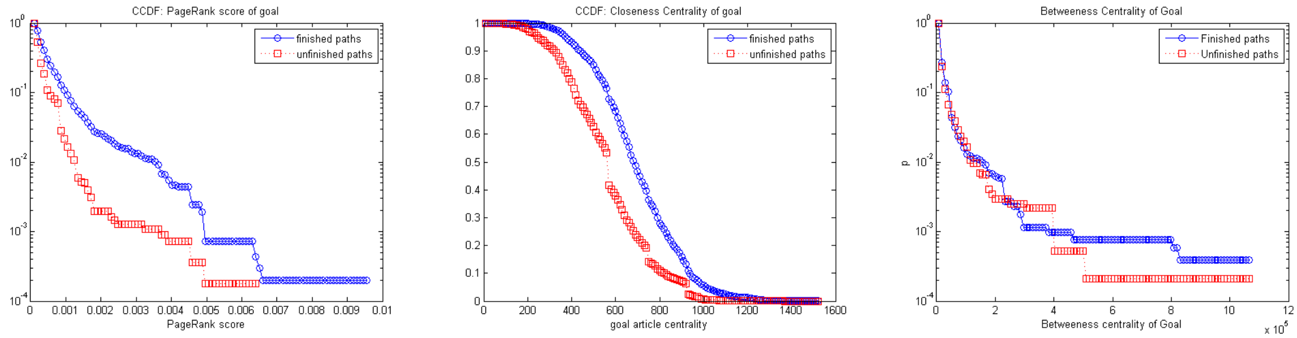
PageRank The PageRank of an article is the most significant measure in our case, as shown in Table 2. In all cases, the centrality of the start article is insignificant. The difference in goal closeness centrality between finished and unfinished paths is relatively slight (less the 5%), with an average centrality value of 684 for finished paths and 550 for unfinished paths. However, when not all nodes are treated equally, and links from central nodes receive higher regard in the centrality computation, as in PageRank, we encounter a much more significant difference between finished and unfinished paths. The average goal article PageRank score among finished paths is 4.2×10^{-4} and 2.1×10^{-4} among unfinished paths.

Article centrality Lower goal article centrality in unfinished paths fits our intuition, as one would expect a mission with a less central goal to be inherently harder. On the flip side, the centrality of the start article does not seem to have an effect on whether a search will end or not, as both metrics of centrality of the start article display almost identical distributions for both finished and unfinished searches. This can imply that the goal article may have a more significant impact than the start article on defining the mission's "a priori hardness". Figure 3 shows that goal centrality, according to all three definitions, exhibits a heavy-tail distribution, suggesting a very small minority of missions have a very central goal. The majority of missions have a more "isolated" goal.

Lucrative ratio of start goal Figure 4 shows that the lucrative ratio of the start node (the fraction of the start node's neighbors that are closer to the goal than the start target itself) is significantly higher for finished paths. This measure indicates an easier start for the finished games.

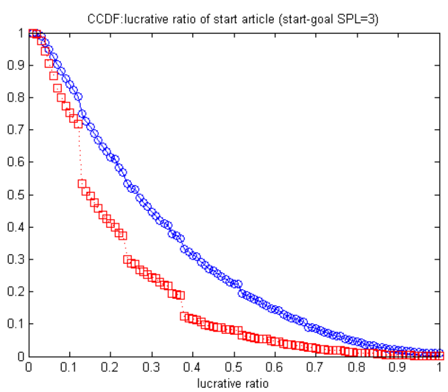
3.c Measurements that depend on user's path

In this section we measure and analyze the properties that depend on the entire path undertaken – the original mission plus the user's choices. Most of the measurements in this category follow the evolution of article properties along the search paths. We perform measurements for four specific missions for which there are many finished and unfinished paths, as well as for all missions combined. Most of the measurements consist of prop-

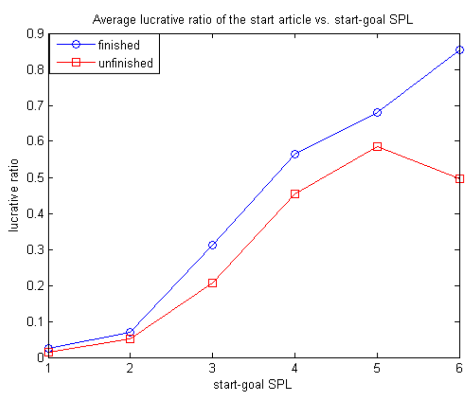


(a) PageRank score (b) closeness centrality (c) betweeness centrality

Figure 3: CCDF of properties of goal article



(a) CCDF of lucrative ratio at start



(b) Avg. lucrative ratio of start node

Figure 4: Lucrative ratio at start

erty evolution along the path. These four missions are listed in Table 3. Table 4 shows a priori hardness criteria for these 4 missions compared with the average finished path and average unfinished path.

Filtering data To facilitate our analysis we restrict ourselves to paths of SPL of 3 clicks between start and goal articles (which is the most common SPL value for both finished and unfinished paths). The game length distribution for both finished and unfinished and both full and effective paths is presented in Figure 5, The full path would include both the back-clicks and then return back, where as effective paths do not include articles visited and then gone back to earlier nodes, as defined in [2]. Very short paths (up to a single click) are common among unfinished games but rare among finished ones. This can be explained by an initial lack of interest on the user's side (not even trying), and we disregard such paths in our measurements. When the user gives up after two or fewer clicks, we do not consider that to be meaningful statistic for comparing with other unfinished games. Likewise, paths of a very large effective length show that either the user is abnormally persistent or potentially a spider. Short paths (of up to 8 clicks) are the vast majority of relevant finished and unfinished paths, and are more common for finished games. We focus on this majority of paths and only consider games with effective path of length 4 to 9 articles (3 to 8 clicks) in our analysis of article properties evolution along the path. We also limit ourselves to games of effective path lengths between 4 and 9.

Intuitively, a high number of back-clicks (or undone clicks) may imply a confused or desperate user behavior that may bolster the chances for giving up. However, Figure 5 also shows that the length differences between full and effective paths are similar in finished and unfinished games. This implies that the number of back-clicks or undone clicks has little effect on whether a search will be completed.

Probability the user will give up Given that we are at the i-th article in the path, what is the probability

Label	Start	Goal
TZ	Theatre	Zebra
AV	Asteroid	Viking
PB	Pyramid	Bean
BT	Brain	Telephone

Table 3: Four most frequent missions

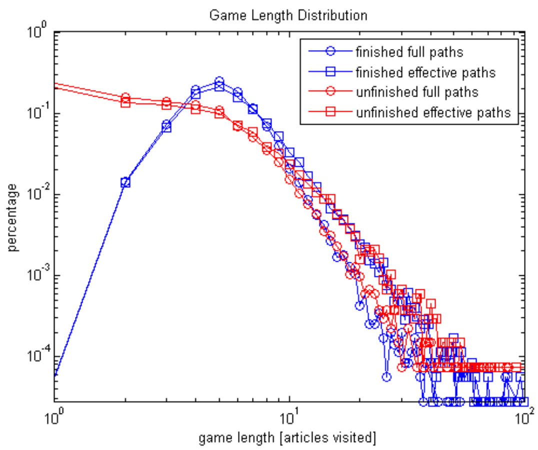
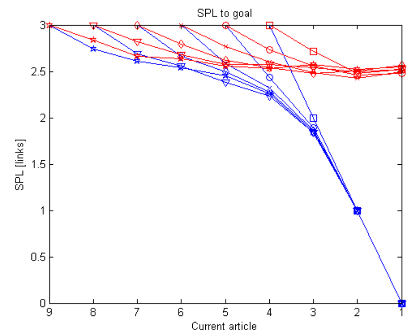
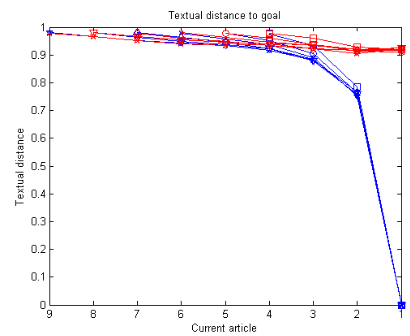


Figure 5: Distribution of game lengths



(a) Evolution of distance



(b) Evolution of textual distance

Figure 6: User progression

that the user will give up? The probability of giving up at the i-th article is given by $Pr(g|i)$:

$$Pr(g|i) = \frac{|\text{unfinished games of length} \geq i|}{|\text{games of length} \geq i|}$$

We see this probability decreasing until a minima for length of 8, and then sharply increasing as the game length increases.

When navigating in a directed graph, there is a chance of falling in a "trap" - a node from which there are no links to the goal. The back-click option may not serve as a savior if the user is not aware of the trap. However, we observed that such a scenario does not actually occur. This is mainly explained by the fact that most of Wikipedia's network is contained within a single giant strongly connected component (SCC) that covers more than 87% of the articles. Since almost all games had both the start and goal articles contained in this SCC, getting out of this SCC and "falling into a trap" is unlikely and does not stand as a major search quitting cause.

User progress The extent of progress the user makes can be measured by how much closer to the goal the user gets with each click. We consider two metrics of distance: 1) SPL to the goal, and 2) textual distance to goal (TF-IDF based, as defined earlier). The distance to the goal at each step is shown in Figure 6b. Each relevant game length (4 to 9 articles) is represented by two curves in these figures - the blue represents finished paths and the red represents unfinished paths.

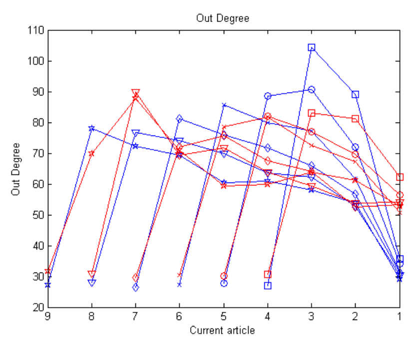
In [2] it is suggested that the strategy used by most users consists of three phases: In the first phase, the user is "getting away" from the start article to a nearby hub. In the last phase, the user is "homing-in" rapidly on the target. The phase between these two is described as a "circling" phase in which the user keeps a relatively constant distance to the goal and fails to make progress. Figure 6b suggests that the user makes progress in a similar manner for both finished and unfinished paths in the first few clicks. However, the user fails to make the transition from the "circling" phase to the "homing-in" phase, and gives up after "circling" within a constant distance from the goal.

The lack of textual distance progress in later phases among unfinished searches is also demonstrated when measuring the textual distance to the next node on the path. While the textual distance between subsequent articles along the path becomes smaller and smaller in early phases (the same way it does in finished searches), in later phases we see this distance becomes almost constant. This suggests again an inefficient "circling" phase before giving up.

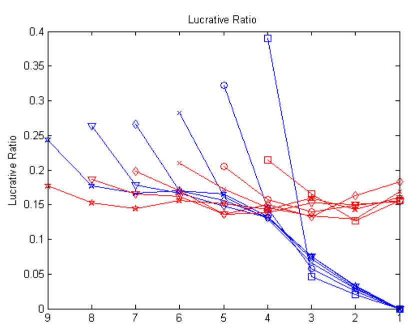
The failure of the user to "home-in" on the target is also supported by evolution of other properties - out-degree, lucrative degree and lucrative ratio, in Figure 7. Notice that the user seems to quit at nodes with an out degree of (roughly) around 60. While this is not a

	TZ	AV	PB	BT	Finished	Unfinished
no. (%) finished paths	404 (54.4%)	444 (59.4%)	264 (37.2%)	426 (54.4%)	34999	-
no. (%) unfinished paths	338 (45.6%)	304 (40.6%)	445 (62.8%)	357 (45.6%)	-	8847
start-goal SPL	3	3	3	3	2.88	3.26
no. of shortest paths	3	3	3	3	2.97	3.29
start-goal textual dist.	0.995	0.973	0.993	0.971	0.976	0.979
PageRank of goal	0.00011	0.00075	0.00012	0.00033	0.0004	0.00019
closeness cent. of goal	556.2	911.5	558.1	739.2	683.9	550
betweenness cent. of goal	19135	41297	4082	7895	15234	32354
lucrative ratio of start	0.118	0.368	0.118	0.222	0.313	0.205
out degree of start	17	19	17	54	27.3	30.4

Table 4: Statistics for the 4 most frequent missions, and overall average for qualified searches.



(a) Evolution of out degree



(b) Evolution of lucrative ratio

Figure 7: Out-degree and lucrative ratio along the path

hub, it is significantly higher than the average degree of a node in the Wikipedia graph (which is around 26, and close to the average degree of a goal), and consistent with the degrees of the nodes in the circling phase. This suggests that even though the user doesn't home-in on the goal, it does not home-in on another wrong node with a small degree. For the frequently-selected mission of BT: the start article (brain) has a high degree (54). In this case the user does not seek a higher degree second node (hub), as the start node itself is almost a hub. The significant difference in the average lucrative ratio between finished and unfinished paths is discussed in 3.2. Specifically, the strategy of choosing a hub in the first click remains, as seen in Figure 7.

Hub quality The *hub quality*, that is the degree of the second article divided by the degree of the maximum-degree neighbor of the start article, is similar in both finished and unfinished. In [2] it is suggested that better information seekers choose better hubs. We see that hub quality seems to have a very minor effect on whether a search is completed (average hub quality for completed searches - 0.49, average hub quality for incomplete searches - 0.43). The four missions show some inconsistent (random) behavior with this respect - i.e. in some missions the hub quality is higher among finished searches, whereas in some missions in is lower., but the differences are always very small.

Evolution of Clustering Coefficient along the path: vaguely an inverse behaviour to that of the average out-degree of the an article along the path (comes from clustering coefficient definition). From looking at the plot for all missions combined, we see that the average clustering coefficient of the article in which the user quit is significantly higher than the coefficients of previous articles along the path or the coefficients of the goals in successful searches. This may suggest that entering a highly clustered part of the graph may encourage quitting. The four most-frequent missions have start articles with a very low clustering coefficient, even though 3 of them also have a low out-degree. The clustering coefficient of an article along the path does not fluctuate

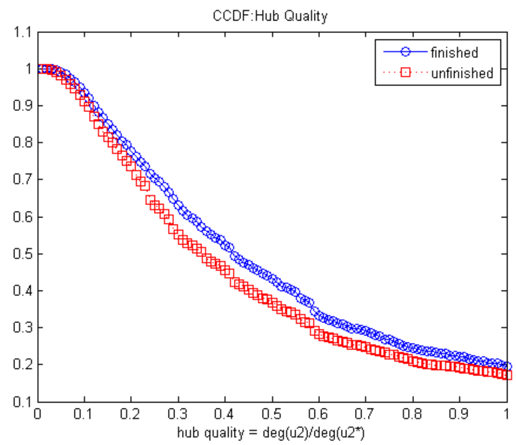


Figure 8: Distribution of hub quality

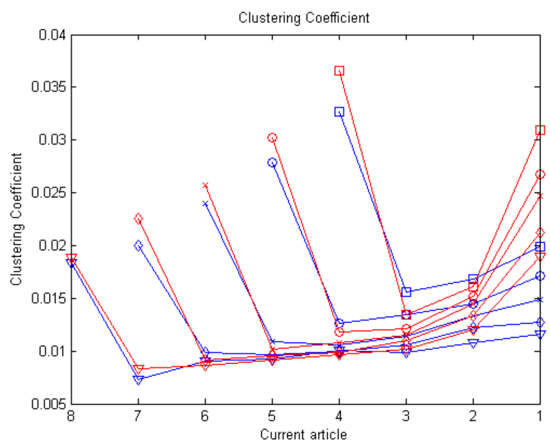


Figure 9: Evolution of Clustering Coefficient

much until later phases, as seen in Firgure 9. This is not the typical case (i.e. the plot of all missions combined), where the start article has a high clustering coefficient that sharply drops when we move to a hub in the first click.

3.d Centrality evolution along the path

We apply the same metrics for centrality along the path as we did for the a priori mission: PageRank, Closeness, Betweenness centrality. Figure 10 shows how these values change along the path. For, PageRank, there are similar trends for both finished and unfinished paths - except for the last node. In unfinished paths, the user seems to give up on a page with a PageRank higher than the goal's PageRank, and similar to previous pages' PageRanks. The trends in PageRank are the same as the trends in the out-degree, and the same reasoning applies here. Closeness centrality, Figure 10b exhibits a similar same behavior as for the PageRank.

The behavior of betweenness centrality, Figure 10b, is similar for finished and unfinished games, even though the centrality seems to be consistently slightly higher for the i-th article on a finished path than for the i-th article on an unfinished path. When fixing the start and goal articles, we see that in most cases betweenness centrality is significantly higher for finished paths. This is perhaps the most distinct feature that distinguishes finished and unfinished paths in earlier phases of the search. While the betweenness centrality of each node along the path seems to be higher for finished paths, the betweenness centrality of the goal (to which the user never reaches in unfinished paths) seems to be much higher in unfinished paths. This, together with the general trend in search of going from low centrality to higher centrality articles, may suggest that failure to move to more central nodes as a contributing factor to search failure.

3.e Back-clicks analysis

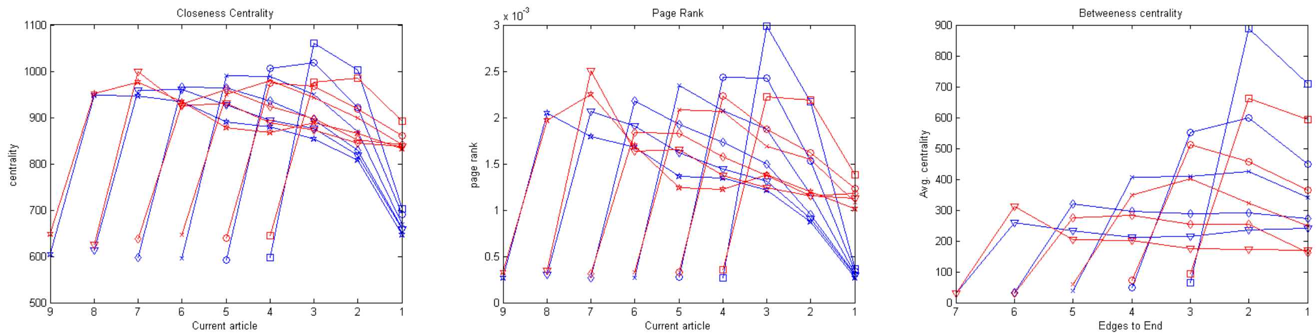
From the game length distribution, in 5, we can see that back-clicks are not common and do not have a significant effect on the game length (the game length distribution for full and effective paths is almost identical). We will call a diversion from the effective path a *tentacle*. A tentacle is not necessarily a chain, but can have any topology. A tentacle is back-tracked using back-clicks and does not affect the effective path.

We took a closer look into back-click patterns, and measured the following:

1. Average number of tentacles, Figure 11 and average length of a tentacle, Figure 12, out of the i-th article as we go along the path. The length of a tentacle out of the i-th article on the path refers to the aggregate length of all the tentacles going out of that article. We see that the number and length of the tentacles (per visited article in the effective path) is very low (less than 1 tentacles per 10 visited articles and roughly 1 back-click per 10 visited articles). This holds for both finished and unfinished games. The only difference is quite obvious - towards the end of a finished game (i.e. homing in on the goal phase), the number of back-clicks drops down sharply.

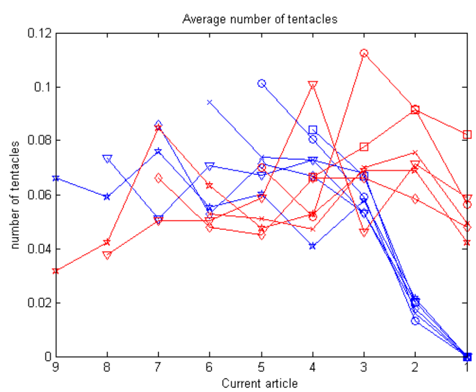
2. The distribution, over all the paths, of the maximal number/length of tentacles out of a single node. Again, we see that the majority of games did not include back-clicks, and that there is no significant with that respect between finished and unfinished games.

While back-clicks can indicate various user moods, possibly frustration, or going back to following a better path), we actually see that they do not have much effect on whether a game is completed.

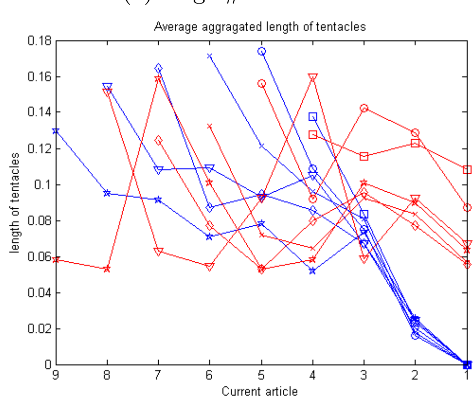


(a) closeness centrality (b) PageRank score (c) betweeness centrality

Figure 10: Evolution of centrality along path.

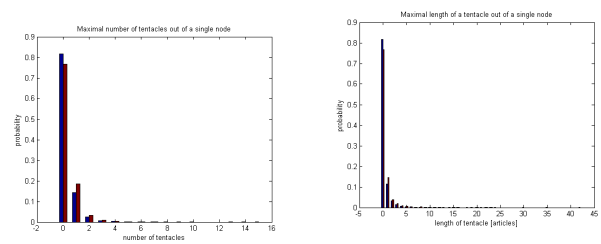


(a) Avg. # of tentacles



(b) Avg. total length of tentacles

Figure 11: Properties of user back-clicks



(a) Max # of tentacles (b) Max length of tentacles

Figure 12: Distribution of tentacles from nodes

Decoys We considered the notion of "decoys" where a decoy is a notion along the traveled path which is "tempting" in some sense but actually take the user farther away from the goal. We tried to see if such decoys would be more common in unfinished paths. More formally we defined a decoy with respect to property p to be an article v such that, for an absolute decoy:

$$(d(v, goal) > d(prev(v), goal)) \bigcap (p(v) > Threshold)$$

And for a "relative decoy," that the new distance is greater than previous distance:

$$(d(v, goal) > d(prev(v), goal)) \bigcap p(v) > p(prev(v))$$

We tried measuring the probability of a decoy as a user makes progress along the path, however we arrived at the same behavior as the probability to just stay away from the target, removing any significance from whether the node is tempting.

4 Further Discussion and Future Work

The user failed on a priori harder paths Most features of the mission itself (a priori hardness) seem to have coherent and sometime significant differences between finished and unfinished searches. A greater shortest path length from the start to target node, a lower lucrative ratio at the start, and a lower PageRank of the goal node are the strongest indicators that the search will lead to the user giving up.

Difference in "homing-in" Measurements that depend on the entire path are more complicated. Most measurements of property evolution along the path are not very helpful, as a significant difference between the behaviour for finished vs. unfinished games only begins to appear in late phases (homing in on the goal for finished games). Furthermore, some characteristics of paths taken (such as back-clicks patterns) seem to not play a significant role in whether the mission will be completed. The most significant property along the path seems to be betweenness centrality of the links.

Features for a learning algorithm The properties for which we have seen differences between finished and unfinished paths could be used as features for a learning algorithm that would try to predict whether a game (or more generally a network search task) would be completed, based on the mission features and on a partial path taken. At every point along the path, there is a probability that the user will "give up." One can apply a machine learning to identify when a user is likely to fail having embarked on a given path, $q =< u1, ..., u_k, >$, and conditional on the target node t. The logistic regression model can be used to classify whether the user will "give up" at a certain point.

Identifying when people are likely to fail A motivator for the project is to identify when people are likely to fail, and to potentially preempt the failure by intervening, or simply prepare for the failure. Alternatively, we may wish to cause an adversary to fail by altering the network. For example, in a terrorist network, we may wish to cause a terrorist to "give up" in their search for a partner in crime, or in finding the materials for building a explosive device. Thus, it would be useful to figure how we can change the network, by severing specific paths or forging new decoy nodes, to increase the failure rate of reaching specific nodes. There are numerous examples of when it would be useful to applying an understanding of when a user will give-up in network search.

5 Acknowledgments

Bob West for creating Wikispeedia and providing invaluable data for this project.

6 References

References

[1] R. West, J. Pineau, and D. Precup. Wikispeedia: An online game for inferring semantic distances between concepts. In IJCAI, 2009

[2] R. West, and J. Leskovec. 2012. Human wayfinding in information networks. In WWW, 2012.

[3] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. PNAS, 102(33):1162311628, 2005.

[4] S. Milgram. The small-world problem. Psychology Today, 2(1):6067, 1967.

[5] D.A. Bader, S. Kintali, K. Madduri, and M. Mihail. Approximating Betweenness Centrality. The 5th Workshop on Algorithms and Models for the Web-Graph (WAW2007). 2007.

A Appendix: Additional data

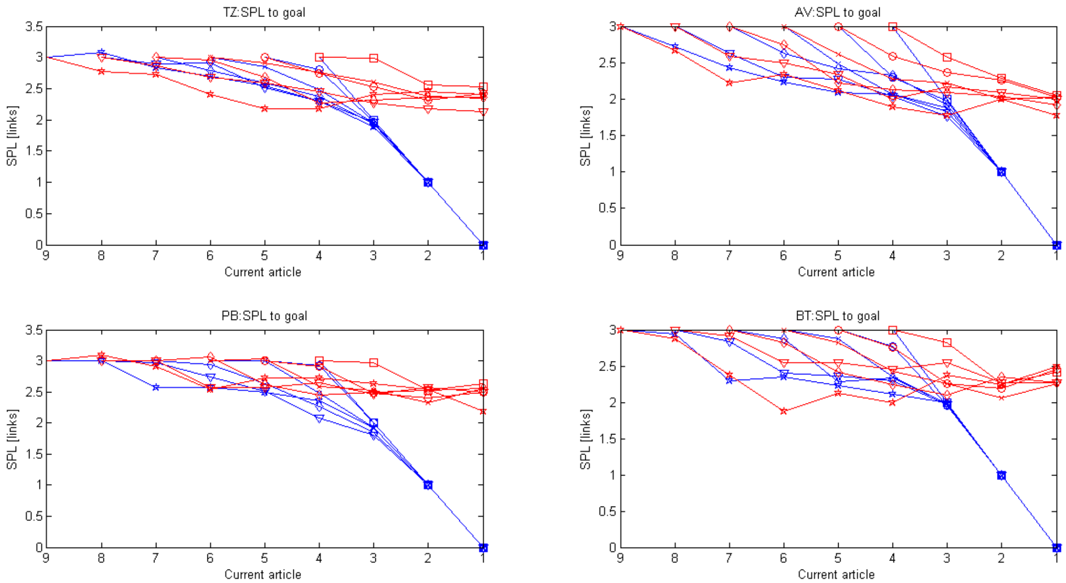


Figure 13: Evolution of distance

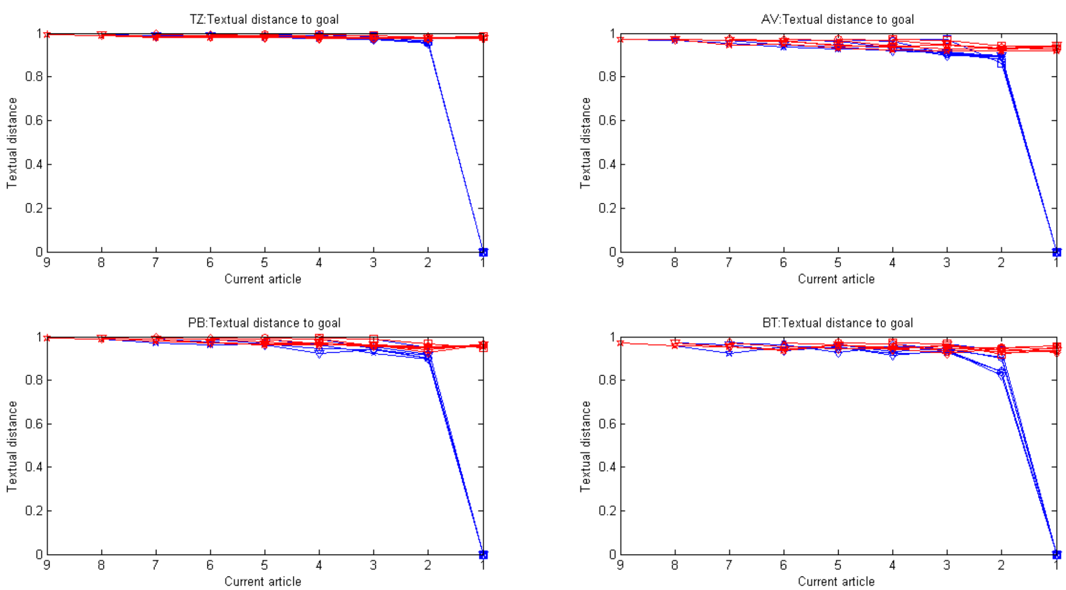


Figure 14: Evolution of textual distance

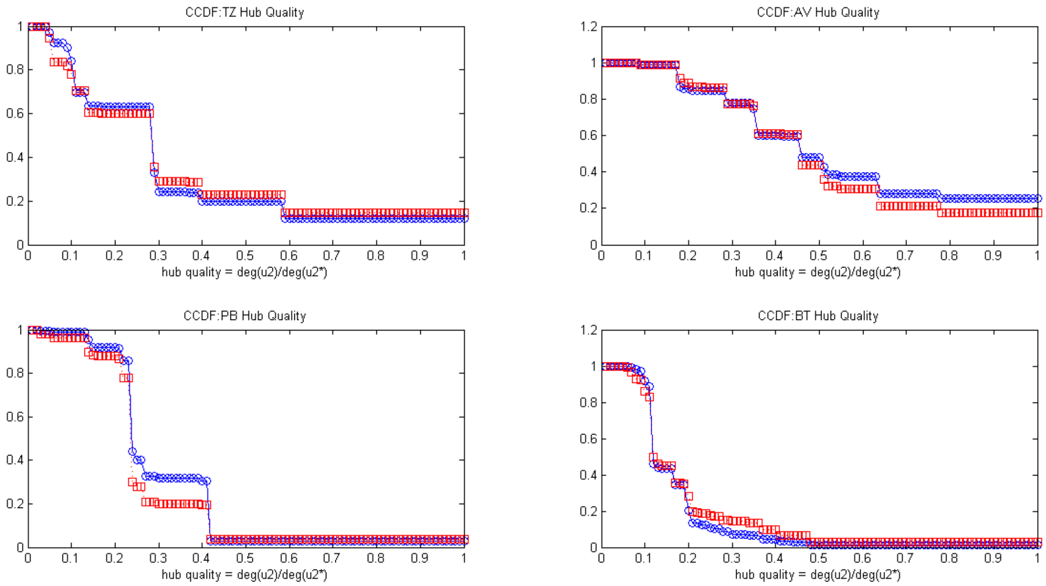


Figure 15: Distribution of hub quality

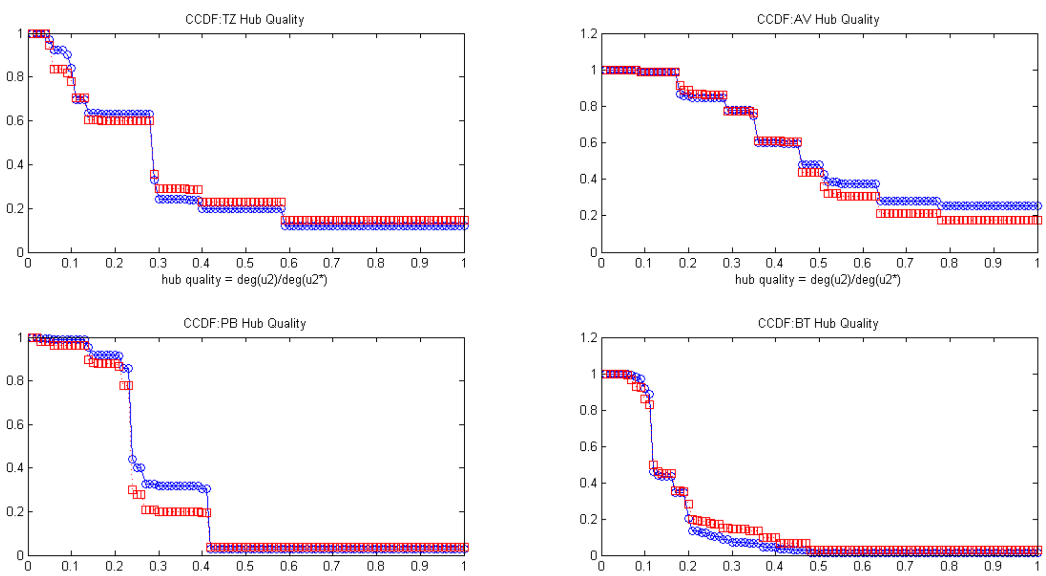
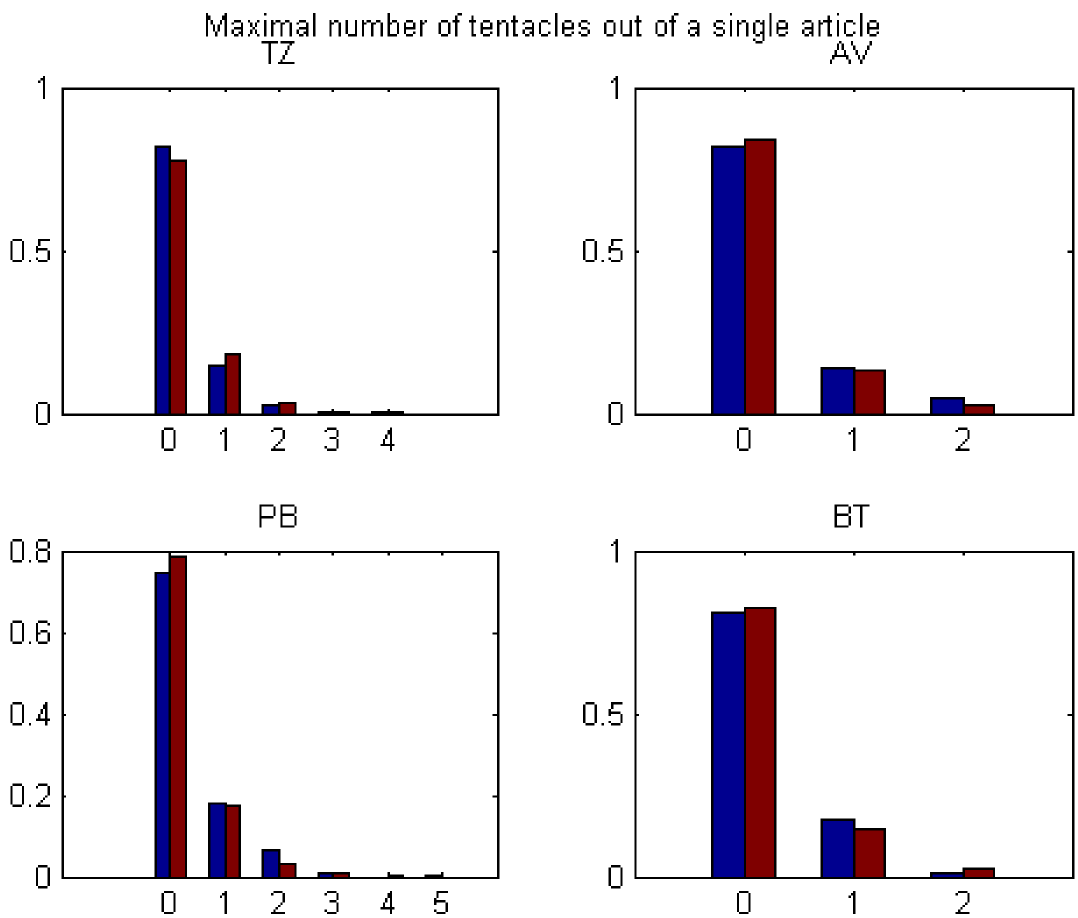
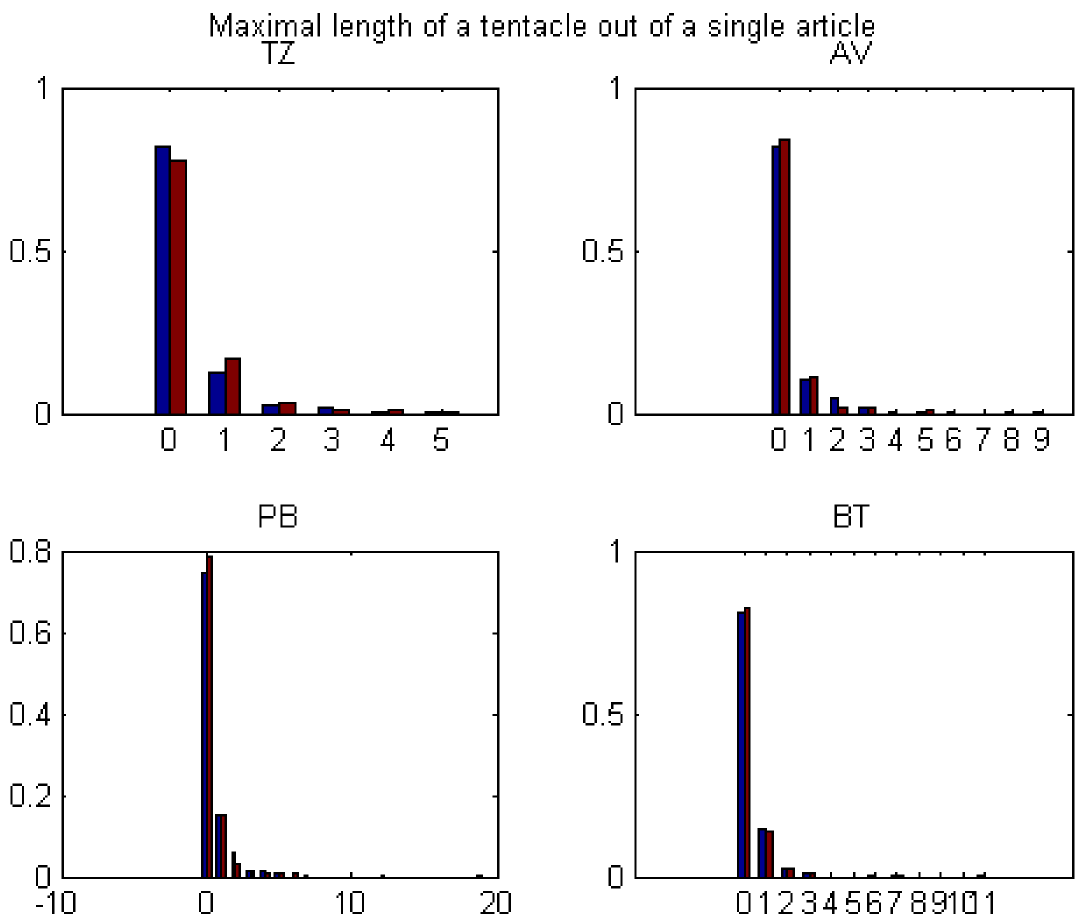


Figure 16: Hub quality along the path



Maximal number of tentacles out of a single article
TZ
AV
PB
BT
1
0.5
0
0 1 2 3 4
0 1 2
0.8
0.6
0.4
0.2
0 1 2 3 4 5


Maximal length of a tentacle out of a single article
TZ
AV
PB
BT
1
0.5
0
0 1 2 3 4 5
0 1 2 3 4 5 6 7 8 9
0.8
0.6
0.4
0.2
-10
10
20
0 1 2 3 4 5 6 7 8 9 10 11