Supporting the Encouragement of Forum Participation

Aashna Garg
Stanford University
aashna94@stanford.edu

Andreas Paepcke
Stanford University
paepcke@cs.stanford.edu

ABSTRACT
Abstract goes here

1. INTRODUCTION
2. INTRODUCTION
3. FROM POSTS TO CONNECTION GRAPH
Social networks are most simply modeled by considering each participant as a node, and interactions initiated by participants as out-directed links. In this case all nodes are of one type, and links are unidirectional. Multiple interaction initiations by one person are captured by weighting the corresponding outgoing links. Many graph analysis tools operate on models of this type, and this is the approach we chose.

However, other strategies exist to cover different goals. For example, [1] additionally consider linkages between forum post topics to include communication content in the model. When networks operate on particular platforms, such as underground forums, which include private 'buddy' connections, such facilities may need to be modeled [2].

For the purpose of identifying candidate time points for encouraging online conversation participation our chosen model suffices. We are not in this work considering additional measures, such as content quality, for which a richer model would be required.

Many measures are used to quantify various aspects of social graphs [**?**]. Not all are meaningful in the context of education-related forum interactions. We focus here on two measures: *out-degree*, and *page rank*. Figure 1 illustrates.

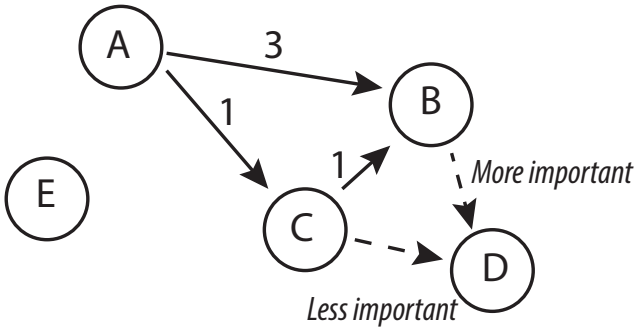

Figure 1: Example social graph induced by forum posts.

Nodes A, B, C, D, and E represent students. The link from A to B is marked with the number 3, because A commented three times on one of more of B's posts. The number of outgoing links is the node's *out-degree*. For example, the *out-degree* of A is 4.

The number of incoming links is called the node's *in-degree*. Node C's in-degree is 1. Node E has no links entering or exiting. The respective student has not participated in the forum.

Analogous to Web pages, each node can be assigned a *page rank*. The intuition in this context is that student S_1's presence in the forum is more 'important' than student S_2's if the node representing S_1 has higher page rank than the node that represents S_2. In our context the intuition behind page rank is that a node N is more important (has higher page rank) the more other important nodes comment on N's posts. Imagine a scenario in which student S_1 posts an interesting question, to which many students comment with their opinion, creating a long thread. The node representing S_1 would experience an increase of its page rank with every incoming comment. Node B in Figure **??** is an example for this situation. Its in-degree is 4. If B were to comment on one of D's posts, then D's page rank would increase more than if the low-page-rank node C commented on D.

In terms of evaluating a student's participation in the forum, a high page rank, and high out-degree are positive. Low values are less positive. We chose these two values because of their relatively straight-forward meanings when applied to forum posts, and for their relevance to our goal of identifying potential intervention times.

Some of the fifteen other measures we computed, such as *betweenness* are meaningful for forum scenarios as well, but their usefulness depends on one's analysis goals. For example, [3] include several of those measures for the purpose of prediction analysis. For evaluation contribution quality the contents of posts would need to be considered: students who persistently post irrelevant contents contribute less positively to the forum than constructively participating students. However, for our purposes the two measures of page rank and out-degree provide strong enough signals.

4. ANALYSIS PROCEDURE
We computed our chosen social graph measures for 39 offerings of seven residential courses, and for one MOOC. Table 1 summarizes these data sources. We recomputed the mea-

sures for every week throughout the quarter of each offering. In order to establish values against which to compare each student's measures during those weekly checkpoints we each time computed the two measures for (*i*) the average of the ten students who overall contributed the maximum number of posts, and (*ii*) for the median number of contributions at each checkpoint.

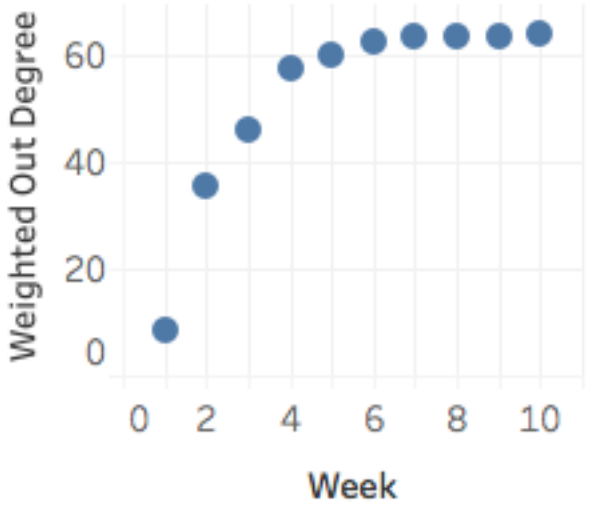

Figure 2: A MOOC forum post comparison: Women's Global Health.

5. CONCLUSION AND FUTURE WORK

Here is an example chart of the correct size.

- Consider student post content quality
- Consider consistency: contribute throughout course
- Consider influence on others
- Draw instructor attention to dense topic clusters, which might indicate confusion, or student excitement to harness.

6. REFERENCES

[1] T. Anwar and M. Abulaish. Modeling a web forum ecosystem into an enriched social graph. In M. Atzmueller, A. Chin, D. Helic, and A. Hotho, editors, *Ubiquitous Social Media Analysis: Third International Workshops, MUSE 2012, Bristol, UK, September 24, 2012, and MSM 2012, Milwaukee, WI, USA, June 25, 2012, Revised Selected Papers*, pages 152–172. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[2] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker. An analysis of underground forums. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, pages 71–80, New York, NY, USA, 2011. ACM.

[3] D. Yang, T. Sinha, D. Adamson, and C. P. Rosé. Turn on, tune in, drop out: Anticipating student dropouts in massive open online courses. In *Proceedings of the 2013 NIPS Data-driven education workshop*, volume 11, page 14, 2013.

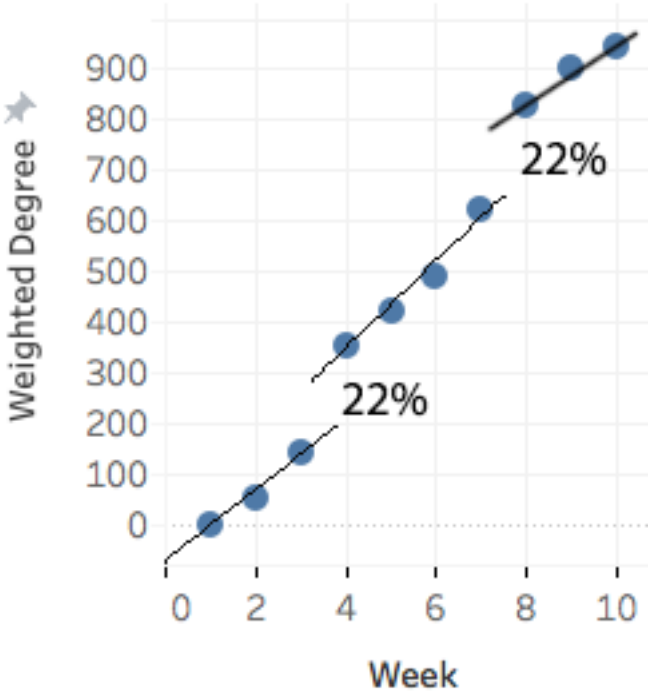

Figure 3: Wrong numbers, just an example. See /tmp/forumPromptsTableauChartsSample.twb

Topic	Offerings	Num-Students	Num-Active-Students	Encouragement
Computational Biology	5	120		
Artificial Intelligence	7	340		
Machine Learning	7	640		
Computer Vision	5	144		
Decision Analysis	4	232		
Audio Signal Processing	6	21		
Political Methodology	5	23		

Table 1: Summary of Examined Courses