

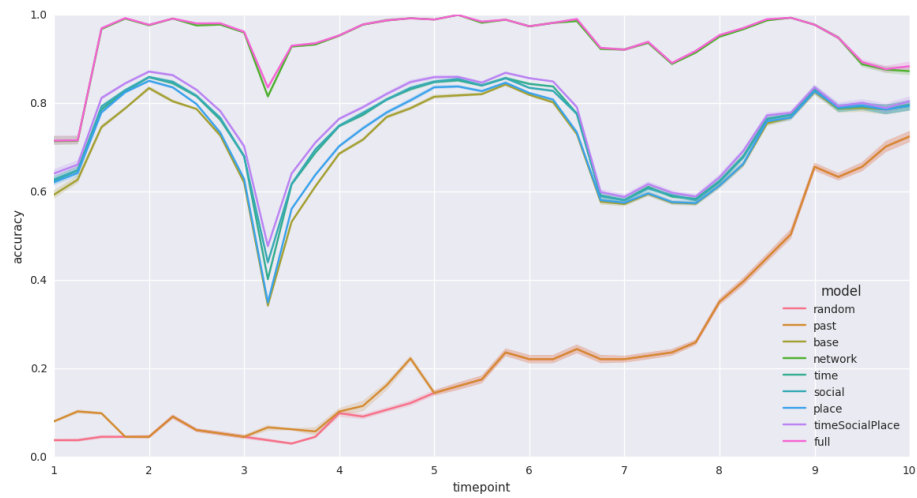
1 Summary

1. I investigated why my model performance was dropping for the later time periods. I found out that the way I select “viable” is introducing a quite significant error (and I didn’t realize that before but almost all link prediction approaches have this problem). I am now not restricting my search space for new nodes to form ties to as the network is actually small enough.
2. I split all ties into four categories (all for the past month):
 - (a) Two people haven’t met
 - (b) Two people have met for less than 15 minutes
 - (c) Two people have met for more than 15 minutes but less than an hour
 - (d) Two people have met for longer than an hour
3. 1) in conjunction with 2) now leads to quite different results. In short, context DOES NOT matter (in contrast to the literature and my behavioral observation), or at least it doesn’t help in predicting who will meet whom (although the model performs quite well in itself now. It is actually really working well now).
4. I found an error in how I calculate the average turnover of ties between time periods. The change in ties is almost 20% per week. I do not think I can call my network a friendship network.
5. I in addition constructed three different types of networks:
 - (a) Social - Those people have to meet outside of university, and are somewhat comparable to your “friends.”
 - (b) University - Those that meet on campus and are your colleagues
 - (c) All - Everybody you meet regardless of where and when
6. I also split the data into more intervals to get more test points. Training and prediction time for an interaction between two nodes stays the same with two months though. I am not exactly sure what the term here is.
7. I added two other baseline scenarios: Random (which is a bit of a misnomer but is basically always guessing the dominant class) and past (which has as its only input whether the class of friendship two nodes had during the previous month).
8. Why it matters? Place in the literature
 - (a) Context does not matter for the formation of social interactions (or at least it doesn’t add much). Contrast to existing literature on place and friendship formation

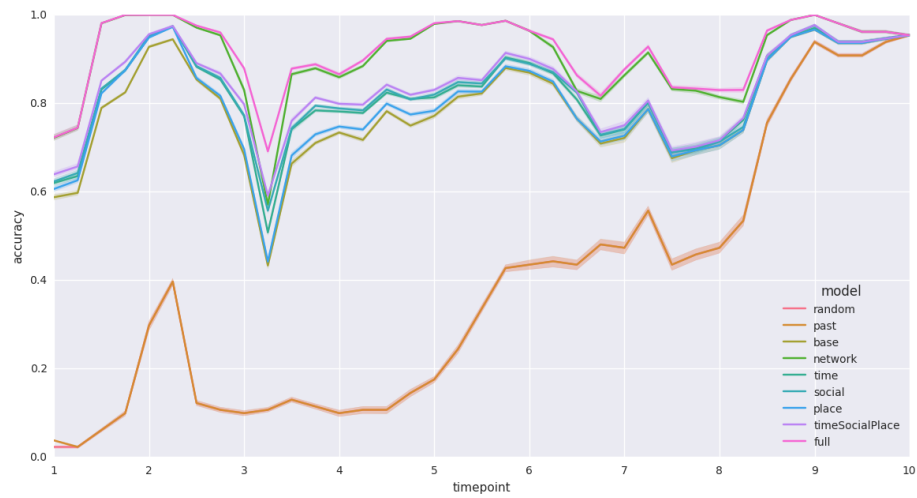
- (b) I haven't seen anybody doing a prediction of social interactions using a link prediction framework (this is a tentative point though for the moment. I want to go back and make sure).
- (c) and last but not least I haven't seen anybody taking space, time, and social "context" into account simultaneously.

2 Model scores

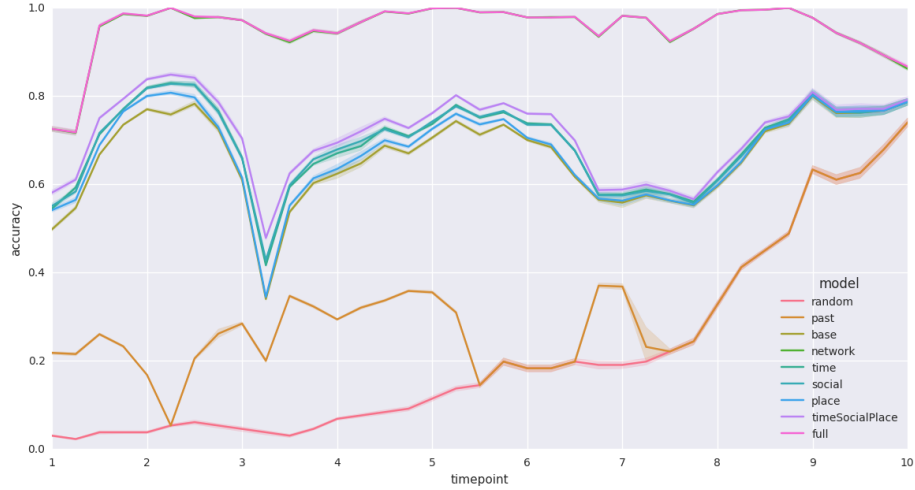
2.1 Social



2.2 University

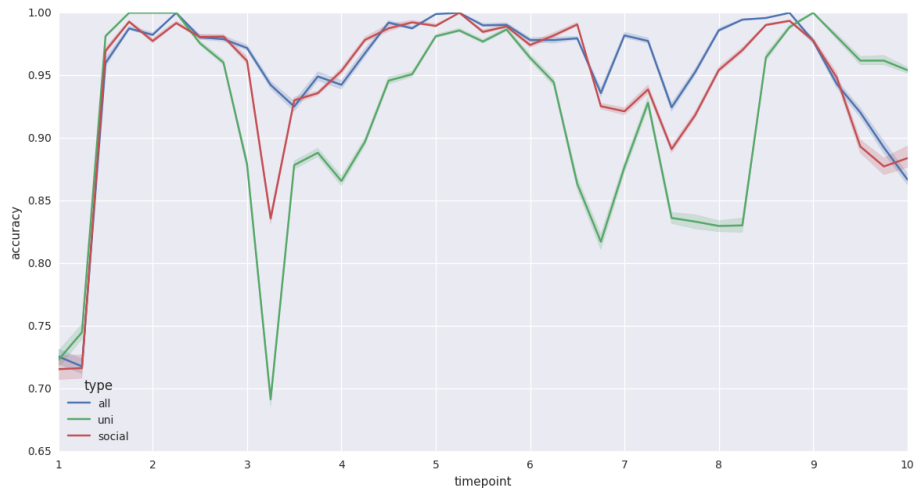


2.3 All



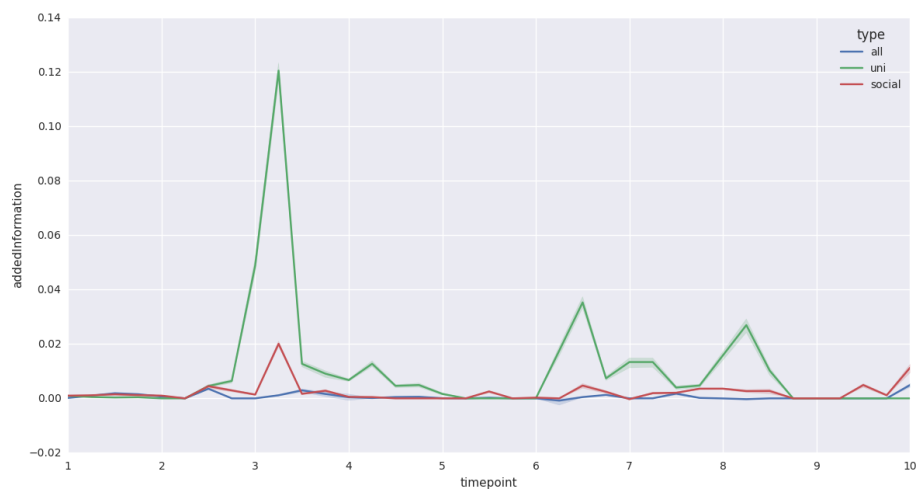
As you can see in basically all models the added information you gain by knowing where and when two people have met is almost negligible. This is in contrast to the current literature that suggests it should play an important role (and I thought so too).

2.4 Comparison of model scores



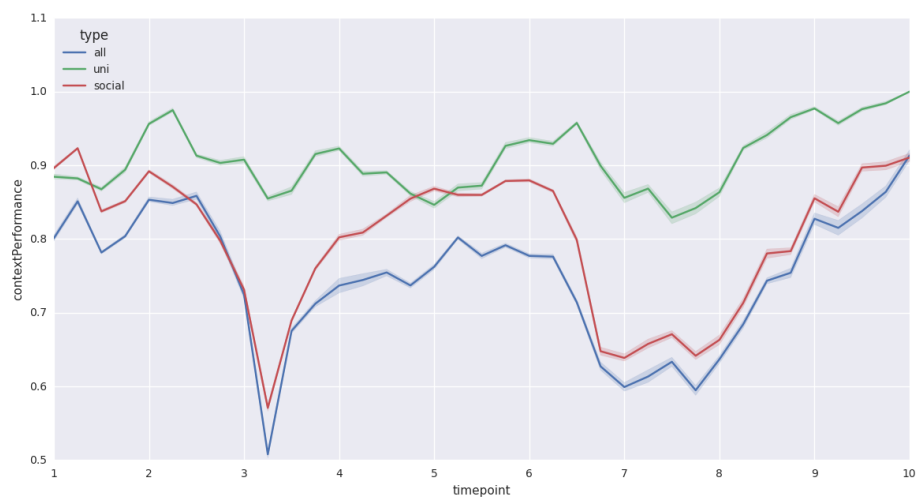
The first drop in performance for all models is Christmas, the second big drop in accuracy is the exam period at the end of the year (and the weeks preceding it to be exact).

2.5 How much does our prediction increase when we add context information?



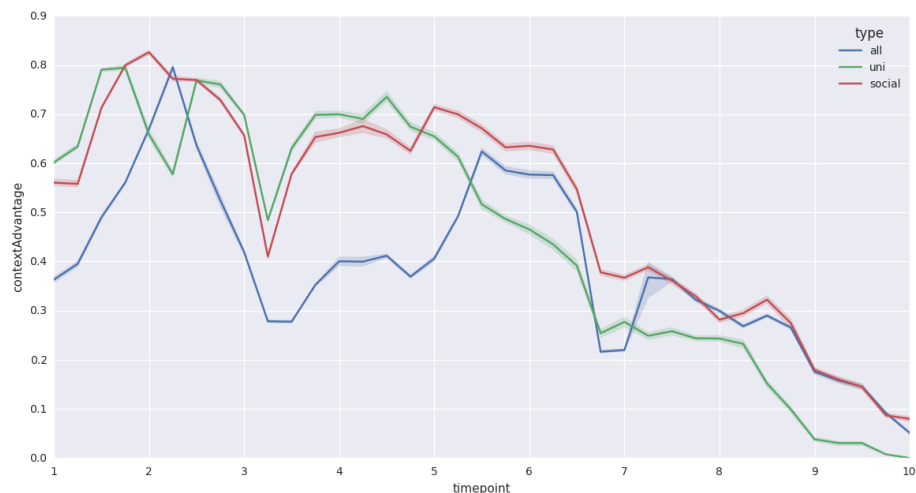
As already pointed out, information about when, whom, and where you have met, does not add much information

2.6 Relative performance of the “context”-only model compared to the full model



However, the “context” only model still compares relatively well to the best prediction we can get using all available information.

2.7 How does the “context” model perform relative to the baseline?



One potential step forward for why context might be useful is if you want to predict future ties, but do not have access to or simply do not know the whole network. In this particular scenario context allows you to improve your prediction significantly (technically there is still the variable “spatial triadic closure” in this model, but I doubt that the results would change much, but this is on my to-do list).