Modeling Memory Imprints Induced by Interactions in Social Networks

James Flamino[1], Ross DeVito[2], Omar Lizardo[3], and Boleslaw K. Szymanski[1]

[1] Rensselaer Polytechnic Institute, Troy NY 12180, USA
[2] University of California San Diego, CA 92093, USA
[3] University of California Los Angeles, CA 90095, USA

Abstract. Memory imprints of the significance of relationships are constantly evolving over time. They are boosted up by the interactions between those involved in the relationship, and decay between such events, causing the relationships to evolve. But despite the importance of the evolution of relationships in social networks, few research papers explore how interactions in those networks over extended periods of time correlate with the memory imprints of relationship importance. In this paper, we represent memory imprints adapting a well-known cognitive science memory model. Using two unique longitudinal datasets, we fit the model's parameters to maximize agreement of the memory imprints of relationship strengths of a node predicted from call detail records with the ground truth limited list of relationships of this node ordered by their strength. We find that this model, trained on one population, performs well on predictions not only within this population but also on different populations, hinting at the universality of memory imprints of social interactions among unrelated individuals. This paper lays the foundation for exploring how the cognitive modeling of memory imprinting in a population with healthy subjects can be used as an unobtrusive tool to identify individuals suffering from memory malfunctions, such as Alzheimer's.

Keywords: Social network · Cognitive modeling · Memory model

1 Introduction

Interactions between members of social networks drive the evolution of such networks. Yet, the underlying mechanics of mapping of the interaction patterns to the evolution of relationships, and consequently to the memory imprints of the network members, are not clear. A primary reason is the paucity of data that provides comprehensive information on both interactions (such as Call Detail Records) and relationship statuses of members (such as ego network surveys). The past works that use such data focused instead on models using machine learning, natural language processing, or more "black box" approaches [7, 5, 13, 6]. While these approaches often accurately predict the surface-level correlations between interactions and members' memory imprints, they lack any analytical,

explanatory value. An individual's memory recall reflects their perception of the impact of interactions, which each leave an imprint that evolves over time. But how can such a longitudinal phenomena in social networks be directly modeled in analytical form? In a recent paper [8], Michalski et al. implement a cognitive model of memory recall (called CogSNet) and evaluate its efficacy at "nominating" significant relationships between individuals from dyadic communication records. This research provides insight into how patterns of communication map to strengths of memory recall of interacting members of the network. However, the authors formulate their model with a parameter which is not a part of the established, well-known memory model, ACT-R [2]. This added parameter represents the threshold value of a memory imprint at which such imprint vanishes from the memory. Yet, unexpectedly, the optimal values of this parameter are high and are unlikely to correspond to any real human memory processes. In this paper, we adapt CogSNet, improving its compatibility with ACT-R. CogSNet uses three parameters for representing memory imprints. Our adapted model, the Memory Imprint Model (MIM), inherits two CogSNet parameters based on ACT-R model that are applied to the entire population. But we replace the third CogSnet parameter with a new one defining for each individual a threshold for the significance of relationships with their partners. This change is important since we discovered that the optimal fit for the MIM parameters are actually universal for populations of peoples with healthy memories. Provided that individuals with malfunctioning memory will recall past interactions and relationships differently than healthy ones, we hypothesize that MIM could be an unobtrusive tool for diagnosing memory malfunction in humans (we elaborate more on this later).

In the rest of the paper, section 2 outlines the datasets used to evaluate MIM. Then, in section 3, we explain how MIM differs from CogSNet, and how we map interactions between individuals into an evolving memory imprint. We compare performance of MIM to a standard Hawkes model and three baseline models. Section 4 discusses our evaluation process in detail. In section 5, we assess the universality of model parameters and find that parameters learned on a population of healthy subjects are valid for other populations of similar individuals. Finally, section 6 summarizes all our results, discusses implications of our results and our hypothesis, and proposes subsequent future research.

2   Datasets

Testing the mechanics of memory imprints arising from interactions in social networks requires a longitudinal study that provides communication data of individuals within a social network, as well as subjective ego network survey data. The former is used to provide dyadic events that can be fitted by the model to simulate the memory imprint left by interactions between the two involved individuals. The latter is used to provide actual empirical measures of significance between the two involved individuals as recorded by the individuals themselves. Unfortunately, data that fits these requirements is scarce. Of the past works

mentioned above, most of those having data that meets these requirements are cross-sectional, relying on testing their models on snapshots of social networks but not exploring the evolution of those networks over time. One exception is the already mentioned CogSNet model that uses date collected in NetSense study [11]. The data were voluntarily collected from randomly selected students entering the University of Notre Dame in 2011. From Fall 2011 to Spring 2013, the participating students had Call Detail Records of their phones continuously recorded. In addition, the participants were required to fill out ego network surveys at the end of each semester. The NetSense study followed 196 students at its peak. We also use data from the followup NetHealth [10] study that involved students arriving in Notre Dame in 2015 and leaving in 2019 using the same methodology as used by NetSense. NetHealth followed 594 students at its peak.

2.1   Communication Records

The communication records for both studies were formatted to conform to the standard CDR format that include for each call a timestamp, sender, receiver, type, and length. While there are auxiliary digital communication channels contained in NetSense (such as WhatsApp), these are sparse, so we keep only the text and phone calls, which provided us $7,465,776$ events. We filter similarly for NetHealth, extracting $41,677,368$ calls and text messages generated by the participants of the study.

2.2   Ego Network Surveys

Each study contains ego network surveys that were collected once a semester to complement the communication records. These surveys were prefaced with a question asking the survey-taker (the ego) to list individuals (the alters) with whom they spend a significant amount of time communicating or interacting. Each ego can list up to 20 alters including people not participating in the study. Thus, the connections of an ego to their alters can involve a variety of relationship types such as fellow students, parents, coworkers, and romantic partners. The ego was subsequently asked about the history of contact with these alters and their shared interests and activities. Importantly, the surveys also asked the ego to subjectively rate similarity and closeness with the alters. The NetSense study has four ego network surveys over the four semesters, and the study participants listed on average 15.9 people per survey. With NetHealth, the average was just 8.4 for the eight ego network surveys over eight semesters.

2.3   Quantifying Relationship Significance

To model the significance of interactions using memory imprinting, we need an empirically-based measure of the significance of relationships that acts as ground truth. Previously, in [8], the authors establish their ground truth as the list of alters nominated by the ego. According to the authors: "These nominations

are based on students' perception of the corresponding relations as one of the top twenty most interacting peers in the surveys administered to participants." However, while the set of the alters themselves can be considered significant to the ego as a whole, there is additional information that we can use from the surveys that provide even greater details on the significance of individual alters, producing a finer measure of significance from which we can more thoroughly test the capabilities of the models in this paper.

In the previous section, we mentioned that there are follow-up questions that are asked about the alters that are listed by the ego. The answers to these questions are mostly selected from a set of modalities that grade the magnitude of an attribute. For example, when asking about an ego's perceived closeness with a listed alter, the choices include "Especially close", "Close", "Less than close", or "Distant". There are also a few questions that require open answers in the form of a rational number. In this paper, we focus on a subset of the inputs from the surveys: closeness, history duration, subjective similarity, emotional significance (for NetSense surveys only), and perceived communication frequency (for NetHealth surveys only). With these question sets, we convert the list of alters into a ranked list, where the higher the rank of an alter, the more significant they are to the ego.

To determine an ordering from the mixed inputs without assuming the importance of any one input over another, we use a pairwise comparative tournament selection process. Consider one ego network survey. Every alter listed by this ego is compared on each question against all of the other listed alters. An alter whose assigned answer to a question is considered "closer to the ego" than that of their counterpart is awarded a point. If the answers of both competitors are of the same value, they are both awarded a point. These points are aggregated across all the questions and then a ranked list is created by ordering everyone by their score in descending order. If there is a tie in aggregate score, the duration attribute is used to break the tie, as it is the only considered input with a rational number answer. After all comparisons are complete for that ego network survey, we are left with our rankings for the top-k significant relations between the ego and their listed alters. The higher the alter is on the list, the more significant relations this alter has with the ego. We run this process for all egos for every survey, yielding our ground truth for this paper.

Given a model and CDR data between an ego and an individual, who interacted with this ego, we fit our model to predict not only if this individual should be listed on the ego survey list, but also what rank, if any this, individual should have on this list. This allows us to assess the model's ability to determine finer levels of significance from the memory imprint from events, not just if the considered relationship has perceived significance.

3   Models

In this section, we outline the models we use to fit ego-alter CDR to nominate alters and predict their ranking in our top-k ground truth lists.

3.1   Hawkes Model

The Hawkes process is a natural choice for simulating the significant relations between two individuals as determined by communication, since the frequency and volume of the communication can often be the key in determining the significance of the relationship. Furthermore, past work directly used this method to understand relationship dynamics [9]. Here, we use a univariate Hawkes process with an exponential kernel for alter nomination and ranking, as defined by the Eq. 1.

$$\lambda(t) = \lambda_0(t) + \sum_{t_i < t} \beta e^{-\beta(t - t_i)} \tag{1}$$

where λ is the Hawkes event intensity at time t (the event time at which the function is being evaluated), t_i is a set of event times before the event time t, and β is the decay rate (which in this scenario represents the decay rate of the memory imprint of an event). We set the arrival rate of immigrant events $\lambda_0(t)$ to 0. Given this function, at survey time t, we consider the communication history of all individuals who have communicated with the ego (as at this point we do not know who is an alter). We feed this data into the Hawkes process up to time t, producing a corresponding λ value for each candidate alter. We then sort these candidate alters in order of descending magnitude of their associated λ value, producing an ordered list of who is significant to the ego according to Hawkes event intensity. We discuss tuning the Hawkes model parameters in subsection 4.1.

3.2   Memory Imprint Model

Our next model, the Memory Imprint Model (MIM), is a variant of the CogSNet model introduced in [8]. Unlike Hawkes, CogSNet uses a recursive piecewise function (cf. Eq. 3) and produces a signal that is normalized to the range $[0, 1]$. We made several adjustments to CogSNet to improve its compatibility with the ACT-R model, and to reduce its computational complexity. First, we define a MIM recall function:

$$r_{ij}(t) = w_{ij}(t_{ij}^b) e^{-(t - t_{ij}^b)\frac{\ln(2)}{L}} \tag{2}$$

where $t_{i,j}^b$ is the time of the last communication event between members (or nodes) i and j in the considered social network before the evaluation time t, and L is the parameter that controls the half-life of node i's memory imprint of their interactions with j. Given this, we define the strength of the MIM memory imprints of interactions between i and j to be

$$w_{ij}(t) = s_{ij}(t) + \begin{cases} r_{ij}(t)(1 - s_{ij}(t)) & \text{when } t > t_{ij} \\ \mu + (1 - \mu)r_{ij}(t)(1 - s_{ij}(t)) & \text{otherwise} \end{cases} \tag{3}$$

where μ is the MIM second parameter that defines the memory imprint strength between i and j. We list alters j's of ego i at each survey time t in the

descending order of their $w_{i,j}(t)$. The length of the list is limited to 20, to mimic the limit for the list length in surveys, and by the threshold defined by the third MIM parameter θ, which limits eligible for listing alters to those whose value of $w_{i,j}(t)$ exceeds θ. Ultimately, MIM introduced here is simpler than CogSNet, and fully compliant with ACT-R. We discuss tuning the MIM parameters in subsection 4.1.

3.3   Baseline Models

Previous research has often used the frequency of communication to predict significance of relations [5, 13, 7, 4, 6]. Following this established methodology, we create the frequency model in the baseline class that calculates frequency (this model's value that is used to nominate significant alters and rank them) by dividing the number of communication events between two individuals by the elapsed time since they first communicated at the timestep for which it is being evaluated. In addition, we create a recency model that uses the elapsed time since the last communication between two people as an inversely related estimate of frequency of contact, as is done in [5]. This produces a recency value that we use to nominate significant relationships and generate the corresponding rankings. Finally, we create a random model that just randomly orders all individuals that the target ego has communicated with, prior to the considered time, into an arbitrary ranking.

4   Results

In this section we detail our evaluation process and present the results for our models with interpretations on their performances.

4.1   Evaluation Process

Given our two datasets are longitudinal, each with their own sets of egos, list of alters, and ego-alter CDR, we designed a comprehensive evaluation process that provides training data for the models to fit their parameters and testing data for the models to test their fit. In the process, we avoid any overlap between training and testing or accessing in learning the future CDR or ground truth data.

Initially, for a given dataset, we use the standard three-fold cross validation to shuffle and then equally split the involved egos into three mutually exclusive groups. Then, for each fold, let N_s be the number of semesters within the target dataset. Within a fold, we separate the training and testing data into N_s subsets, split by semesterly survey time. At each survey time, the training and testing subset only includes the surveys and communication data from before that time (therefore the later subsets contain the earlier subsets). We stagger the communication training data, such that the models only fit their parameters with communication data up to the survey time that precedes the survey time

against which the models are being tested. This design ensures that the models are not able to use any of the training surveys that are at the same time as the current target testing surveys. Subsequently, we start with the second semester for rankings in each fold. After models have fit their parameters with the current training subset and have produced rankings for the current testing subset, the process is then advanced one subset forward, allowing more training and testing data to be released with each following survey time.

Note that the training and testing data are kept completely separate at all times. Testing data, even from previous semesters, is never used to fit the models' parameters, and test fold ground truth from surveys is only used to score a model's rankings. Once all predictions for a fold are complete, we compute the fold's overall score: the weighted average ranking prediction accuracy of the predicted rankings across all available surveys for testing egos.

For ranking accuracy when comparing rankings generated by a model against the top-k ground truth, we use Ranked Biased Overlap (RBO) [12]. We chose it because of its use of an indefinite rank similarity measure for the similarity of two ranked lists. Moreover, RBO handles items being present only in one ranking, weights the agreement of top ranking items higher than that of lower ranking items, and works with any given ranking length. We set the RBO's weight parameter p to 0.98.

The average of the RBO scores for all available testing ego surveys is weighted by the length of the corresponding top-k ground truth lists, accentuating accurate predictions of larger lists. The final score is computed as the average fold score over all folds. This setup allows our models to show if they are able to continuously predict rankings over extended periods of time.

When fitting the models using training data, we use a proven, established framework for optimized parameter fitting: Optuna [1], which uses an efficient approach to parameter sweeping and critical parameter identification when tuning models. Optuna is used to tune β for Hawkes, and L, μ, and θ for MIM each time a new chunk of training data is released during the evaluation process.

4.2   Model Performance

We test the performance of our models in our evaluation process for both Net-Sense and NetHealth, with the results being presented in Table 1. The random model sets the lowest bar for performance, followed by recency and the frequency. The tunable models, Hawkes and MIM, outperform the baselines to a statistically significant degree. Interestingly, we found that Hawkes outperforms MIM slightly, but not on a level that is statistically significant. The minor bump in performance is due to the fact that Hawkes' single parameter allows for faster and more accurate tuning of the model through Optuna. Furthermore, the additional parameters of MIM allow for a greater range of error, especially when fitting the complexities of memory imprint through CDR. Regardless, the results of MIM and Hawkes are promising, considering the fact that these models are not only nominating significant relationships for the fit of their interactions, but are also assigning accurate levels of significance of each relationship.

Table 1. Average RBO scores (variances) for NetSense and NetHealth.

Model	RBO Scores	
	NetSense	**NetHealth**
Random	0.03022 (0.002)	0.03299 (0.004)
Recency	0.28295 (0.029)	0.27508 (0.033)
Frequency	0.29526 (0.030)	0.29526 (0.030)
MIM	0.36058 (0.031)	0.35833 (0.036)
Hawkes	0.36497 (0.031)	0.37117 (0.038)

5   Cross-Dataset Modeling

In this section, we evaluate the generalizability of our models by using Hawkes and MIM tuned on data from one study to nominate and rank alters in the other. Successful cross-dataset ranking performance serves as an indicator that a model is picking up on transferable cognitive behavior that should be universal across populations. In this comparison, we need to consider the different sizes of NetSense and NetHealth datasets. Tuning models on NetHealth to test on NetSense will give them an overt advantage over the times series models because of the size of the training data compared to the testing data. As mentioned in section 2, NetHealth has 594 egos, while NetSense only has 196. Therefore, we randomly remove 6 egos from the NetHealth dataset and then evenly split the remaining egos into three subsets, resulting in three subgroups of 196 egos, equal to the number of available egos in NetSense. Then to equalize the number of social tie rankings, we repeat the splitting twice, with the egos first keeping their first four semesters of data, and then the last four semesters of data. This results in six NetHealth subgroups that can be used independently for training and testing with NetSense.

Here, we evaluate only the memory models, since they have parameters that can be tuned, and since the two datasets were collected during entirely different times, we do not need to worry about staggering the training and testing process as done previously. When training on NetHealth to test on NetSense, we tune the models on one of the six NetHealth subgroups and then test it against all of the NetSense data, then repeat for all subgroups. When training on NetSense to test on NetHealth, we tune the models just once on NetSense and then test on each NetHealth subgroup. For both approaches, the average of all RBOs weighted by top-k list length gives us the subgroup scores, the average of which creates the final score. The results of this evaluation process is shown in Table 2.

Importantly, the models have similar prediction performance between the previous results and in the cross-dataset modeling results. This suggests that these tuned models are picking up on memory imprint mechanics that are universal across datasets. The universality of these mechanics implies that while communication patterns and overall behavior between populations may change, the way individuals processes and perceives them in memory over time does not. Our models are trained on subjects with healthy memories. Hence, we can

Table 2. The dataset on the left of the arrow is the dataset on which the models were exclusively trained. The dataset on the right of the arrow is the dataset the models were tested on after training. The score is average RBO with the variance in parentheses.

Model	Training Dataset → Testing Dataset	
	NetHealth → NetSense	NetSense → NetHealth
MIM	0.35970 (0.0005)	0.33063 (0.0021)
Hawkes	0.36293 (0.0006)	0.35071 (0.0016)

conclude that they are effective not only at modeling memory imprints within single social networks, but also on social networks with healthy members.

6   Conclusions

Our results show that the Memory Imprint Model and the Hawkes model are capable of effectively capturing memory imprints of interactions. This process not only allows us to identify significant relationships from interactions over time, but also allows us to assign levels of significance to the relationship compared to others, giving us a finer degree of explanatory power when evaluating connections between people in social networks. Importantly, our results on cross-dataset modeling shows that these models, when fit on the memory imprint dynamics of one population, can be used to accurately nominate and rank significant relationships of the other population. Based on these results, we hypothesize the MIM and Hawkes models capture the dynamics of a healthy memory, as we are studying the cognitive dynamics of populations of young adults starting college. Consequently, the participants whose mechanics of memory imprint we are modeling will be, generally, in their prime health-wise. Hence, the accuracy of the results for cross-dataset modeling confirms the two models' efficacy in representing the mechanics for the memories of young, healthy adults.

In future work, we will attempt to prove or disprove our hypothesis that the models and the process we have designed for evaluating said models can be considered a viable standard for representing memory imprint of interactions for healthy memories that can ultimately be used as a comparative tool for diagnosing malfunctioning memory imprint patterns in humans. Specifically, if we tune our models to nominate significant relationships using a population of healthy adults and then test it against a population where some individuals have malfunctioning memory (i.e., Alzheimer's), the model performance will be negatively affected, as the memory imprints of the ailing individuals would not match the memory imprints of the healthy population. The easily and unobtrusively obtainable differences may indicate the type and level of advancements of illnesses, like Alzheimer's or dementia, better than the current indirect method of diagnosis of the memory aliments [3].

References

1. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: A next-generation hyperparameter optimization framework. In: Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2019)
2. Anderson, J.R., Bothell, D., Lebiere, C., Matessa, M.: An integrated theory of list memory. Journal of Memory and Language **38**(4), 341–380 (1998)
3. Bayat, S., Babulal, G.M., Schindler, S.E., Fagan, A.M., Morris, J.C., Mihailidis, A., Roe, C.M.: Gps driving: a digital biomarker for preclinical alzheimer disease. Alzheimer's Research & Therapy **13**(1), 1–9 (2021)
4. Bulut, E., Szymanski, B.K.: Exploiting friendship relations for efficient routing in mobile social networks. IEEE Transactions on Parallel and Distributed Systems **23**(12), 2254–2265 (2012)
5. Conti, M., Passarella, A., Pezzoni, F.: A model for the generation of social network graphs. In: 2011 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks. pp. 1–6. IEEE (2011)
6. Flamino, J., DeVito, R., Szymanski, B.K., Lizardo, O.: A machine learning approach to predicting continuous tie strengths. arXiv preprint arXiv:2101.09417 (2021)
7. Gilbert, E., Karahalios, K.: Predicting tie strength with social media. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 211–220 (2009)
8. Michalski, R., Szymanski, B.K., Kazienko, P., Lebiere, C., Lizardo, O., Kulisiewicz, M.: Social networks through the prism of cognition. Complexity **2021** (2021)
9. Nurek, M., Michalski, R., Rizoiu, M.A.: Hawkes-modeled telecommunication patterns reveal relationship dynamics and personality traits. arXiv preprint arXiv:2009.02032 (2020)
10. Purta, R., Mattingly, S., Song, L., Lizardo, O., Hachen, D., Poellabauer, C., Striegel, A.: Experiences measuring sleep and physical activity patterns across a large college cohort with fitbits. In: Proceedings of the 2016 ACM international symposium on wearable computers. pp. 28–35 (2016)
11. Striegel, A., Liu, S., Meng, L., Poellabauer, C., Hachen, D., Lizardo, O.: Lessons learned from the netsense smartphone study. ACM SIGCOMM Computer Communication Review **43**(4), 51–56 (2013)
12. Webber, W., Moffat, A., Zobel, J.: A similarity measure for indefinite rankings. ACM Transactions on Information Systems (TOIS) **28**(4), 1–38 (2010)
13. Wiese, J., Min, J.K., Hong, J.I., Zimmerman, J.: "you never call, you never write" call and sms logs do not always indicate tie strength. In: Proceedings of the 18th ACM conference on computer supported cooperative work & social computing. pp. 765–774 (2015)