Love All, Trust a Few: Link Prediction for Trust and Psycho-social Factors in MMOs

Muhammad Aurangzeb Ahmad[1], Zoheb Borbora[1], Jaideep Srivastava[1], and Noshir Contractor[2]

[1] Department of Computer Science, University of Minnesota, Minneapolis, MN, USA
{mahmad,zborbora,srivastav}@cs.umn.edu
[2] School of Communication, Northwestern University, Evanston, IL, USA
nosh@northwestern.edu

Abstract. Massively Multiplayer Online Games (MMOGs) where millions of people can interact with one another have been described as mirrors of human societies and offer excellent venues to analyze human behavior at both the psychological as well as the social level. Within the context of predictive analysis (link prediction as a classification task) in MMOGs, the connection between psycho-sociological theories of communication networks. A mapping of how various elements of trust and other social interactions (mentoring, adversarial relationship, trade) relate to prediction tasks is also established. Results from classification experiments indicate that social environments affect prediction tasks in cooperative vs. adversarial environments in MMOGs and the implications of these results for generalizability of link prediction algorithms is also analyzed.

Keywords: Trust in social networks, Prediction and Psycho-Social Theories, Adversarial environments, MMOs.

1 Introduction

One of the seminal events of the last decade has been the explosion of myriad arrays of various form of social media which generate gigabytes of data every hour and thus provide an unprecedented opportunity to analyze human behavior on a massive scale. Mainly because of this data revolution it is now possible to not just build better theories regarding human behavior but also move from a descriptive analysis of social data to a predictive analysis. One issue which is usually coterminous with predictive modeling is that it is often the case that the models do not explain the psychological and social reasons behind *why* the model is successful in predictive analysis and thus essentially a black box. We consider these issues in the context of the link prediction problem.

While the problem of link prediction has been studied before in a number of contexts in social networks, we note that this problem has not been addressed with respect to the role of social science theories to explain the efficacy of featuresets in prediction tasks. One step in that direction is work by Ahmad et al [4]

S.J. Yang, A.M. Greenberg, and M. Endsley (Eds.): SBP 2012, LNCS 7227, pp. 123–130, 2012.

who try to incorporate Monge and Contractor's Multi-Theoretical Multi-Level framework [10] in the link prediction tasks. We take their work one step further by linking the feature space to theory space and additionally describe how the results of prediction tasks can be interpreted in terms of social science theories.

2 Background

The link prediction problem consists of a family of prediction problems which may range from predicting the formation [9], breakage [11], change of links to recurrence in the edge formation [12]. The link prediction problem was first described by Liben-Nowell and Kleinberg [9] and the Inter-Network Link Prediction Problem was first described by Ahmad et al [4] who also proposed a social science theory based approach to address that problems. In a follow up work Borbora et al [5] explored the problem of efficacy of feature space associated with link prediction to determine a robust set of features for link prediction.

Model based explanations for predictive modeling can be divided into three main types: (i) Explanations regarding how the algorithm works (ii) Explanations regarding how the model explains the phenomenon, such explanations are usually absent from black box models e.g, Neural Nets. (iii) In social, psychological and cognitive domains explanations that link the model to motivations that can be ascribed to intentional agents (people) or groups of such agents (society). In recent years there has been a move towards linking prediction algorithms, models and feature spaces to explanations in terms of social and psychological theories when these involve social phenomenon. That is mainly because an explanation agnostic model would not gain much currency in the social science domain where the primary goal is to not just study these phenomenon but also provide explanations with respect to why things happen. Borbora et al [6] thus note the distinction between theory driven and data driven models and how one can inform the other in creating better predictive models.

3 A Psycho-Social Framework for Link Prediction

The MTML framework [10] describes the creation, maintenance and development of linkages in social networks in organizational and inter-organizational contexts and links together various theories in the sociology literature which also harkens to psychological motivations regarding why people form relationships with one another. We refer the reader to the text by Monge and Contractor [10] for a detailed descritionb of the theories of communication networks. We take these theories as starting point in feature set construction and also the partition of the feature space based on the appropriate features. We use the feature-set scheme used by Hasan et al [7] and modified by Ahmad et al [4] as our starting point but we expand it to include additional features which are more appropriate for a larger social space. They divide their feature space into three sets of features as follows: (i) Proximity Features (ii) Aggregated Features (iii) Topological Features. We note that this classification scheme is based on how the featuresets are

Table 1. Mapping Between Feature-sets and Theories in the MTML Framework (i)

	Self-Intr.	Cognition	Evolution	Exchange	Contag.	Homo.	Prox.
Ascribed							
Human Gender						X	
Avatar Gender						X	
Avatar Race						X	
Country						X	X
Σ Human Age						X	
ΣAvatar Age						X	
Human Age Diff.						X	
Avatar Age Diff.						X	
ΣJoining Age	X						X
Joining Age Diff.	X						X
Acquired							
Char Class Ind.	X						
Char Level Sum		X					
Char Level Diff.		X					
Guild Indicator		X			X		X
Guild Rank Sum	X	X					X
Guild Rank Diff.	X	X					

constructed with minimal or no regard to their relationship to motivations with respect to why people form links. We expand their scheme and extend the set of features and first divide them based on how they are described in the sociology literature. Thus the Proximity features can be mapped to Ascribed (attributes based on some intrinsic node characteristics) and Acquired characteristics (node characteristics which can change in time) [2]. The topological features mostly map onto the social neighborhood based characteristics. Additionally we introduce a new class of characteristics i.e., trans-social characteristics which span multiple social networks and are defined as indicator functions i.e., if the node n belongs to the network A then the value of the function is 1. The mappings between the theories and the featuresets is given in Tables 1 and 2.

4 Experiments

We use a dataset from a massively multiplayer online game called EverQuest II (EQ2) where players can interact with one another in multiple ways and there are many avenues of socialization so that it is possible to construct multiple coextensive social networks between them. To check how well the classification tasks do in different social environments, we use data from two different servers or social environments in EQ2. One of the servers (called guk) represents a cooperative or neutral environment, called Player vs. Environment (PvE). The other server (called Nagafen) represents an adversarial environment, Player vs. Player (PvP). Our main motivation behind using different social environments

Table 2. Mapping Between Feature-sets and Theories in the MTML Framework (ii)

	Self-Intr.	Cognition	Evolution	Exchange	Contag.	Homo.	Prox.
Social Neighborhood							
Degree Cent. Diff	X		X				
Betwn. Cent. Diff	X		X				
Σ Degree			X			X	X
Degree Diff.			X			X	X
Shortest distance					X		
Σ Clustering Ind.	X		X	X			
Common Neighbors				X			
Salton Index	X	X	X				
Jaccard Index	X	X	X				
Sorensen Index	X	X	X				
Adar-Adamic Index	X	X	X				
Resource Alloc. Index	X	X	X				
Trans-Social							
Trust Link		X	X	X			
Mentor Link		X	X	X			
Trade Link		X	X	X			
Group Link		X	X	X			
Combat Link		X	X				X

was that the social relationships would form differently in the two networks and thus that would be reflected in the efficacy of the prediction algorithms even though the same feature sets are used in the feature space. The network characteristics of these networks are given in Table 3 where NCC refers to the number of connected components. We use a binary classification approach for link prediction as proposed by Hasan et al [7] for link prediction within and across social networks [4]. The dataset is divided into training period and test period. For each of the tasks 60,000 instances are prepared for prediction. A positive example is when the edge does not exist in the training period but exists in the test period. In the case of the negative example the edge does not exist in either periods. We used a standard set of classifiers (Naive Bayes, Bayes Net, KNN, SVM, JRip, J48, Adaboost) for our experiments and report for best results for each classification task.

The results of the experiments for the two networks are given in Table 4. The source network refers to the network which is used to construct the training examples and the destination network is the network for whom the prediction has to be made and is from the test period. The main thing to note here is that while the results for many of the prediction tasks remain more or less the same, in a subset of the cases there is a marked difference between the results that we get for the adversarial environment vs. the cooperative environment. The cases which are markedly different for the two environments are highlighted in a different color in Table 4. Thus consider the prediction results for the mentoring network, as noted in previous work [4] and [5] the prediction performance for

Table 3. Network Statistics for all the networks used

Type	Network	Nodes	Edges	Diameter	NCC
PvE	Trust	15,465	23,145	37	1,488
PvP	Trust	13,184	15,945	27	2,237
PvE	Mentor	23,207	93,079	39	316
PvP	Mentor	36,973	187,452	≤ 27	97
PvE	Trade	31,900	1,796,438	≤ 24	11
PvP	Trade	49,132	2,142,832	≤ 24	20
PvP	Combat	59,468	3,767,395	≤ 24	32

the mentoring network is relatively low as compared to the other networks. However the results for the same prediction task but in the adversarial network are much better. This difference can be attributed to the fact that just as the adversarial environment results in greater competition between players who are in opposing teams and thus adversaries, the opposite is also true for people in the same teams i.e., one would expect greater loyalty for players in the same teams in adversarial environments as compared to people who are in cooperative environments. This results in overall better prediction results for the mentoring network prediction task. A similar difference is noted for the prediction tasks from mentoring to trade as well as from the trade to the mentoring network. Again, in both the cases the results for the adversarial environment are better as compared to the cooperative environment. The main take away from these observations is that the mentoring network is a better predictor for links in the trade network and vice versa in the adversarial environment as compared to the cooperative environment and for the same reasons. While the mentoring and trade results are commutative in this case, this is not true for the prediction tasks for the trade and trust networks i.e., there is a marked difference in the prediction results for trade to trust and not vice versa for the two environments. The main reason for this result is that a trade edge has a low transaction cost associated with it as compared to a trust edge which has a high cost associated with it. Thus a trust relationship is likely to have a corresponding trade relationship associated with it but not vice versa. Theories of co-evolution [10] would imply that in cooperative environments neutral and positive interactions (trade and trust respectively) are likely to percolate from one dimension to another but this is less likely in adversarial environments which explains the results.

We note that given the nature of the two environments the Combat network is not present in the PvE server. Additionally we have access to another network in the PvE environment, called the grouping network, which was not extracted for the PvP environment at the time of these experiments. The group network refers to an ingame network formed by players who group together to complete quests. These are analogous to military missions or other logistical missions in the offline world. The results for the Combat network in the adversarial environment and the results for the Grouping network in the cooperative environment are given in Table 5. Overall the results for the combat network as well as the grouping

Table 4. Results for Link Prediction in Adversarial vs. Cooperative Environments

Networks		Cooperative			Adversarial		
Source	Destination	Precision	Recall	F-Score	Precision	Recall	F-Score
Trust	Trust	0.79	0.69	0.74	0.79	0.66	0.72
Mentor	Mentor	0.63	0.48	0.54	0.77	0.71	0.74
Trade	Trade	0.80	0.78	0.79	0.86	0.85	0.86
Trust	Mentor	0.67	0.43	0.52	0.64	0.47	0.54
Trust	Trade	0.75	0.73	0.74	0.78	0.79	0.78
Mentor	Trust	0.88	0.76	0.82	0.85	0.67	0.75
Mentor	Trade	0.74	0.74	0.74	0.84	0.85	0.84
Trade	Trust	0.89	0.83	0.86	0.88	0.75	0.81
Trade	Mentor	0.67	0.55	0.60	0.81	0.75	0.78

network are quite good even when compared against other prediction tasks. The main exception in this case is again the mentoring network where the prediction results for grouping to mentoring network are not as good as the other prediction results. The main reason for this observation is that while a large number of mentoring instances co-occur with the grouping instances i.e., mentoring occurs in the context of grouping in such cases but the opposite is not necessarily true i.e., grouping usually does not co-occur with mentoring [3].

5 Interpretation and Methodological Issues

We have considered the problem of link prediction in the context of two different social environment and a feature space mapped onto different social science theories. Our main motivation for using two different social environments is to highlight the hazards of generalization without considering the social environments associated with the prediction task. Thus consider previous results reported by Hasan et al [7], Ahmad et al [4] and Borbora [5] using similar techniques and link prediction tasks in general, the generalizability of the feature space is assumed without the social context. Theories in the social sciences e.g., the MTML framework [10] imply that social networks in different social environments evolve differently which is in turn reflected in their network structure. The differences in the network structures are also likely to affect prediction and this is in line with some of the observations that we made in the results.

There are additional methodological issues with respect to generalizing across MMOG environments. Thus consider the case of modeling of team formation dynamics in the online world by Johnson et al [8] who show that the same generative models can be used to explain guild formation in World of Warcraft and street gangs in Los Angeles. Based on their observations they generalize that there must be some common generative mechanism for team formation in online guilds and offline street gangs. Ahmad et al [1] replicated their results in EQ2 and discovered that the generalization does not carry over to EQ2. More research is required to settle this issue conclusively but both of these cases highlight the fact

Table 5. Results for Link Prediction for the Group and Combat Networks

Networks		Adversarial			Networks		Cooperative		
Source	Destination	Precision	Recall	F-Score	Source	Destination	Precision	Recall	F-Score
Group	Group	0.88	0.90	0.89	Trust	Combat	0.88	0.91	0.89
Trust	Group	0.88	0.90	0.89	Mentor	Combat	0.84	0.85	0.84
Mentor	Group	0.85	0.83	0.84	Trade	Combat	0.88	0.90	0.89
Trade	Group	0.86	0.86	0.86	Combat	Combat	0.89	0.91	0.90
Group	Trust	0.87	0.75	0.80	Combat	Trust	0.88	0.74	0.80
Group	Mentor	0.61	0.47	0.53	Combat	Mentor	0.83	0.78	0.81
Group	Trade	0.81	0.83	0.82	Combat	Trade	0.86	0.88	0.87

that generalizations are unwarranted especially in contexts where social contexts are not taken into account.

6 Conclusion

Predictive analysis, especially classification, is an important aspect of machine learning and while the internal mechanism of most classification algorithms are well understood, a mapping of feature spaces to social and psychological theories is not well understood. In this paper we considered such a mapping and used two datasets representing two social environments in an MMOG. The results showed that for a subset of the prediction tasks the prediction models perform differently using the same feature set. This implies that the network structures associated with the adversarial versus the cooperative environments are different and should inform the selection of features for future work.

Acknowledgment. The research reported herein was supported by the AFRL via Contract No. FA8650-10-C-7010, and the ARL Network Science CTA via BBN TECH/W911NF-09-2-0053. The data used for this research was provided by the SONY corporation. We gratefully acknowledge all our sponsors. The findings presented do not in any way represent, either directly or through implication, the policies of these organizations.

References

1. Ahmad, M.A., Borbora, Z., Shen, C., Srivastava, J., Wiliams, D.: Guilds Play in MMOs: Rethinking Common Group Dynamics Models SofInfo 2011, October 6-8 (2011)
2. Ahmad, M.A., Ahmad, I., Srivastava, J.: Marshall Poole Trust me, I m an Expert: Trust, Homophily and Expertise. In: MMOs IEEE SocialCom 2011, Boston, MA, October 9-11 (2011)
3. Ahmad, M.A., Huffaker, D., Wang, J., Treem, J., Kumar, D., Poole, M.S., Srivastava, J.: The Many Faces of Mentoring in an MMORPGs. In: IEEE Social Computing Workshop on Social Intelligence in Applied Gaming, SocialCom 2010, Minneapolis, MN, USA, August 20-22 (2010)

4. Ahmad, M.A., Borbora, Z., Srivastava, J.: Noshir Contractor Link Prediction Across Multiple. In: Social Networks Domain Driven Data Mining Workshop (DDDM 2010), ICDM 2010 Sydney, Australia (2010)
5. Borbora, Z.H., Ahmad, M.A., Haigh, K.Z., Srivastava, J., Wen, Z.: Exploration of robust features of trust across multiple social networks. In: Fifth IEEE Conference on Self-Adaptive and Self-Organizing Systems Workshops (SASOW 2011), October 3-7. Ann Arbor, Michigan (2011)
6. Borbora, Z., Hsu, K.-W., Srivastava, J., Williams, D.: Churn Prediction in MMORPGs using player motivation theories and ensemble approach. In: Third IEEE International Conference on Social Computing. MIT, Boston (2011)
7. Al Hasan, M., Chaoji, V., Salem, S., Zaki, M.: Link prediction using supervised learning. In: Workshop on Link Counter-terrorism and Security. SIAM (2006)
8. Johnson, N.F., Xu, C., Zhao, Z., Ducheneaut, N., Yee, N., Tita, G., et al.: Human group formation in online guilds and offline gangs driven by a common team dynamic. Physical Review E 79(6), 066117 (2009)
9. Liben-Nowell, D., Kleinberg, J.M.: The link prediction problem for social networks. In: CIKM 2003 (2003)
10. Monge, P., Contractor, N.: Theories of Communication Networks. Oxford University Press, Cambridge (2003)
11. Sharan, U., Neville, J.: Exploiting Time-Varying Relationships in Statistical Relational Models. In: SNA-KDD 2007 (2007)
12. Tylenda, T., Angelova, R., Bedathur, S.: Towards Time-aware Link Prediction in Evolving Social Networks. In: KDD-SNA 2009 (2009)