Chapter 6

Discussions

In this chapter, we elaborate on the main contributions of this thesis. Our primary objective in this thesis was to improve the efficiency and effectiveness of Link Prediction and Influence Maximization in dynamic social networks. So we developed in this thesis a series of principled approaches for both link prediction and influence maximization. Mainly we were interested in addressing the highly dynamic behaviour of social networks and also explored the multiple features responsible for the improvement in link prediction and influence maximization in dynamic social networks.

6.1 Summary and Contributions

We summarize and highlight here the main contributions and findings of the thesis in light of the research objectives stated in Chapter 1. Critical discussion of the entire work will also reveal the advantages and limitations of the proposed models. The following subsections summarize and discuss the contributions, advantages, and limitations of the considered objectives of this thesis.

6.1.1 Link Prediction based Influence Maximization

Our first objective in this thesis was influence maximization in highly dynamic social networks. To achieve this objective, we proposed a link prediction-based influence maximization technique in chapter 3. Here, first, we define a novel Influential Node Tracking problem to maximize the influence spread in an online social network. Then, we propose an LPINT framework for efficient and effective influence maximization in dynamic social networks. The LPINT model consists of two steps: Link Prediction and Seed selection. For link prediction, we use the ctRBM technique, which combines the temporal as well as structural behaviour of the nodes in evolving graphs to predict the upcoming links. Next, to select the efficient seed set, we improved *UBLF* algorithm and proposed *EXCHANGE* algorithm in our LPINT framework. In this chapter, we assume that the active nodes are the ones that have communicated in past N snapshots. We find seed nodes only among these active nodes. Implementing this assumption is novel and makes our goal of influence maximization more efficient and effective.

Finally, through experiments, we show that the proposed framework performs better in terms of influence spread in comparison to the considered baseline techniques on considered datasets. We show the improvement in results in terms of influence spread experimentally and theoretically. In our proposed work, once the behaviour of nodes for making the new links are learned, the prediction of the upcoming snapshot becomes efficient and effective. Efficient seed nodes reduce the number of iteration in the IC model for influence spread and hence take less time for information spread as compared to other considered baseline algorithms.

The limitation of our proposed method includes the overhead of prediction of the upcoming snapshot; however, with the increase of time system learns for efficient prediction. Here, we have not considered the situation where any node behaves randomly, although it is also not considered by the baseline algorithms.

6.1.2 Multifeature Analysis based Link Prediction

Our second objective in this thesis was link prediction in dynamic social networks. To achieve this objective, we proposed the Multifeature Analysis-based Link Prediction technique in chapter 4. In this model, we added mobility, popularity, and similar interests of the nodes as additional factors along with the structure and attributes to predict the network evolution pattern and the upcoming links in the evolving social networks. Here, we use an improved LDA topic model and Hidden Naive Bayesian algorithm to propose a PILHNB model for link prediction in dynamic social networks.

Our proposed model learns the individual nodes' behaviour pattern with time, making the model more consistent, robust, and best suited for noisy networks because it considers each users' importance in the network. Considering the location and popularity feature makes the model more accurate. Using common interest and attribute similarity feature makes the model more effective than the considered baseline methods. However, the graph embedding and graph neural networks based baseline methods consider mainly the structural information of the nodes and their neighbours; they do not consider the content of communication messages or the other factors which we have been considered here.

The experimental results validate that the proposed PILHNB model gives better performance in terms of precision, recall, F1-Measure, and AUROC on almost all the considered datasets compared with other considered baseline methods. By using our proposed model, we can effectively predict links among users of social networks. Our model can learn the user behaviour pattern, which changes over time, and also the pattern of structural changes in the networks. It can be applied to understand the evolution pattern of dynamic networks and can be useful in many applications of link prediction.

The limitation of our proposed model is the consideration of many factors together; it makes the model theoretically complicated and increases the preprocessing overheads for finding different feature vectors.

6.1.3 Context-aware Influencial Nodes Tracking

The third objective in this thesis was the context-aware influential nodes tracking in dynamic social networks. Here our aim is to maximize the spread of information in social networks. To achieve this, we propose a multi-feature analysis-based influential nodes tracking model named MINT algorithm for influence maximization in dynamic social networks given in Chapter 5. In our proposed model, we use the structure of the network, users' topic-of-interest, users location sharing information, and the popularity of nodes in the network to propose a context-aware independent cascade diffusion model for influence spread. We also propose an efficient topic-aware seed selection technique that uses the Topic-aware Influence sub-Graph for finding the seed set for topic-based information spread.

The results of the experimental evaluation prove the efficiency and effectiveness of our proposed model. We can observe that using the interest similarity between users, location, and popularity-based assumptions makes the model more effective because these features are based on the real-life scenario of our society. Using a Topic-aware influence graph for topic-based seed selection makes our model efficient and scalable because this model can be easily implemented on large graphs. Using the Topic-aware Influence sub-Graph for seed selection reduces the time and space complexity for the seed selection process compared to the considered baselines.

However, our proposed model can not give any theoretical guarantee for influence maximization, which is a limitation of our proposed MINT model.