**QUICK POLLING TECHNIQUES FOR NETWORK USING THE FRIENDSHIP PARADOX: "WHAT ARE YOU AND YOUR FRIENDS THINK?"**

*1B.Prithviraj Singh , 2S.Anusharma* *,3 A.Jayaprakash, 4**N.Shiva,* **5***T.Ramakrishna*

# *1,2,3,4bachelor Degree, Cse,* *Jayamukhi Institute Of Technological Sciences Narsampet, Warangal.*

*5Mr T.Ramakrishna, Associate Professor, bachelor Degree, Cse, Jayamukhi Institute Of Technological Sciences Narsampet, Warangal.*

**ABSTRACT:**

The topic of this study is randomly polling a community of people. Using a selected group of people, classical intent polling attempts to predict what will happen of an election between contestants A and B by asking them: "Who will you vote for?" Who do you believe will win, according to expectation polling? What is your estimate of the percentage of wins for A, we question randomly selected people in this paper's new neighbourhood expectation polling (NEP) technique. Therefore, while responding to this question in NEP, sampled individuals will automatically consider their neighbours (as described by the underlying social network graph). As a result, choosing the best collection of samples from the network is crucial to minimising the mean squared error (MSE) of NEP approaches. In order to do this, we provide three NEP algorithms for the following scenarios: Although the social network graph (i) is unknown, random walks (also known as sequential exploration) can be carried out on the graph (ii). Algorithms based on the friendship paradox, a graph theoretic consequence, are presented for both situations. The dependency of the MSE of the algorithms on the network's characteristics is proved theoretically.

To demonstrate how well the algorithms work, numerical results on both genuine and made-up data sets are given.

**INTRODUCTION:**

This study examines randomised polling of a social network with a potentially unidentified structural layout. Who will you vote for? is a question that is asked to uniformly sampled respondents in a classical intent poll in order to predict the outcome of an election between candidates A and B. Who do you believe will win, according to expectation polling? What is your estimate of the percentage of votes for A, we question non-uniformly selected people in this study as part of a unique neighbourhood expectation polling method. The problem is then clearly defined, along with the method to solving it and the work that inspired it.

Consider a community that can be illustrated by an undirected graph G = (V, E), where each of the nodes has a label f(v) that ranges from 0 to 1. A total of |S| (also known as the sample budget) people from this social network can be questioned by a pollster.

Issue description. Estimate

f = |{v ∈ V : f(v) = 1}/|V | (1)

which is the fraction of nodes with label 1 , with a sampling budget |S| |V | for the following cases: • Case 1 - graph G = (V, E) is not known but, the graph can be explored sequentially using a random walk.

Case 2 - graph G = (V, E) is not known but, the set of

nodes V can be uniformly sampled

To solve the aforementioned problem1, we provide a family of polling techniques we term neighbourhood expectation polling (NEP). The question "What is your estimate of the fraction of people with label 1?" is posed to a collection S V of individuals from the social network G = (V, E).

Any person instinctively looks to her neighbours when attempting to estimate an unknown amount about the globe.

As a result, each sampled person s S would supply the percentage of their neighbours N (s), with label 1. In other words, the individual's response to the NEP question would be,

q(s) = |{u ∈ N (s) : f(u) = 1}/ |N (s)|

Then, the average of all the responses P s∈S q(s) |S| is used as the NEP estimate of the fraction ¯f

**RELETED WORK:**

As previously mentioned, in the traditional intent polling2 method, a set S of nodes is created using random sampling with substitutes, and the standard deviation of their tags is then calculated.

is employed as an estimate (henceforth referred to as the intent polling estimate) of the proportion f specified in (1). The fundamental drawback of intent polling is that an additive mistake requires a sample size of O( 1 2) [3]. Our study is inspired by two recently put out approaches that aim to get over this issue in intent polling, specifically "expectation polling" [6] and "social sampling" [3].

First, each sampled person offers a projection of the identity held by the vast majority of the network members in expectation polling [6] (i.e., sampled people respond to the question, "Who do you think will win the election?").

). Then, each sampled person will consider her neighbours and indicate the value that most of them share. Comparing this strategy to intent polling, it is more effective (in terms of sample size) because each sampled person now supplies the presumed answer of a neighbourhood.3,4. Second, in social sampling [3], each sampled person's reaction depends on the labels, qualifications, and sampling probabilities of her neighbours. [3] gives limits for the variances of numerous unbiased estimators for the fraction f using this approach.

Then, each participant in the sample will think about her neighbours and mark the value that the majority of them share. This method is superior than intent polling in terms of the number of samples because each sampled individual now provides the presumptive neighbourhood response.2. Each sampled person's response in social sampling [3] is influenced by the labels, credentials, and sampling probabilities associated with their neighbours. For calculating the percentage f using this method, [3] provides bounds for the variances of several unbiased estimation methods.

G = (V, E) , Undirected graph with set of nodes V and set of edges E

A , Symmetric adjacency matrix of the graph G where

A(u, v) = ( 1, if (u, v) ∈ E 0, otherwise

n , Number of nodes i.e. n = |V |

M , Number of friends i.e. M = 2|E|

N (v) , The set of neighbors of a node v ∈ V as defined by the graph G

d(v) , Degree of node v ∈ V i.e. d(v) = |N (v)|

f(v) , Binary label of node v ∈ V

f¯ , Fraction of nodes with label 1 i.e.

f¯ = |{v ∈ V : f(v) = 1}| |V | q(v) ,

NEP response of node v ∈ V i.e. q(v) = |{u ∈ N (v) :

f(u) = 1}| |N (v)| D , Diagonal matrix with

D(v, v) = d(v) A , Normalized adjacency matrix

A = D− 1 2 AD− 1

Random Variables, Distributions and Related Parameters

X , Uniformly sampled node from set of nodes V

Y , Random friend: uniform sampled end of a uniformly sampled edge from E

Z , Random friend of a random node

P(k) , Degree distribution which gives the probability that a random node X has degree k

q(k) , Neighbor degree distribution that gives the probability that a random friend Y has degree k

e(k, k0 ) , Joint degree distribution that gives the probability that a random edge (U, Y ) will have nodes with degrees d(U) = k, d(Y ) = k 0

σk , Standard deviation of the degree d(X) of a random node X i.e. standard deviation of the degree distribution

σf , Standard deviation of the label f(X) of a random node

X rkk , Neighbor degree correlation coefficient defined in (28) ρkf , Degree-label correlation coefficient defined in (29)

**Polling Estimates and Related Parameters:**

S , Set of the individuals queried by the pollster

|S| , Sampling budget (number of individuals queried by the pollster)

N , Length of Random Walk (for Algorithm 1)

I |S| , Intent polling estimate defined in (3)

T |S| UN , Naive NEP estimate with uniformly sampled nodes defined in (7)

T |S| RW , NEP estimate obtained via proposed Algorithm 1 T |S| F N , NEP estimate obtained via proposed Algorithm 2

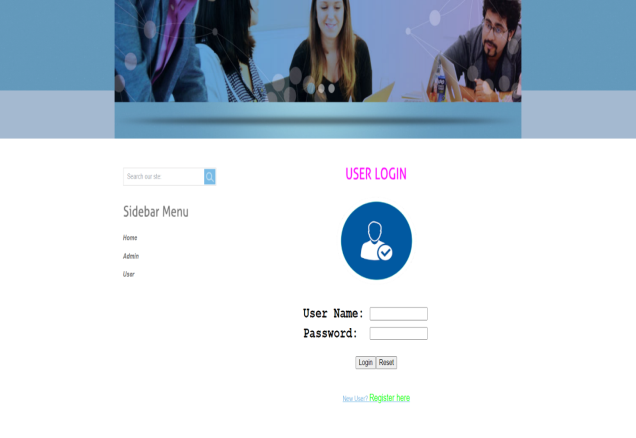
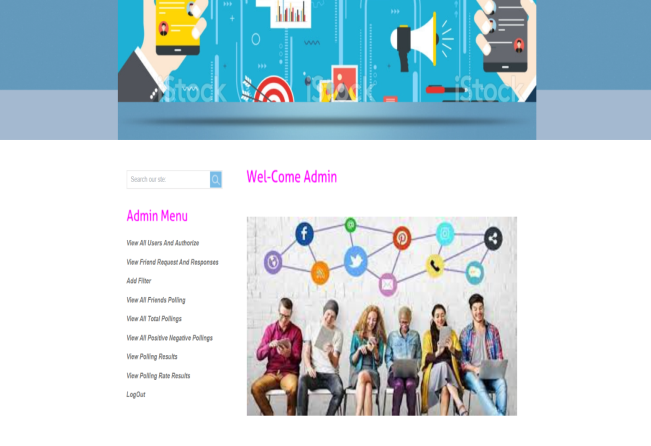
details on a very huge graph. Therefore, NEP may be viewed as a method that asks a question with a finer precision than expectation polling while being more straightforward and logical than social sampling.

The friendship paradox, a type of neural sampling bias seen in undirected networks, is the central concept used in our proposed NEP estimators for cases 1 and 2 (given in the issue statement). Recently, the friendship paradox has drawn attention in a number of network-related applications under the general heading of "how network biases are able to be used effectively for estimation problems?" For instance, [14, [15] demonstrate how the friendship paradox may be used to accurately estimate a heavy-tailed degree distribution, and [16, [17] demonstrate how the friendship paradox can be used to promptly identify a disease epidemic. Our findings in cases 1 and 2 are similarly broadly consistent with this pattern. Friendship paradox has been studied in relation to estimation problems as well as perception biases in social networks [18], [19], [20], information diffusion and opinion formation [21], [22], [23], [24], influence maximisation and stochastic seeding [25], [26], [27], node properties other than degree [28], [29], [30], and directed social networks [18], [28], [31].

evenly, and then selecting one end of it via a fair coin toss). A more natural comparison, however, would be between the degree d(X) of a random person X and the degree d(Z) of a random buddy Z of a random person. This is accomplished through [32]'s development of the crucial refinement of the friendship dilemma.

**RESULT:**





**CONCLUSION :**

In order to estimate the percentage of nodes in a network that have a specific characteristic (represented by a binary label), this study studied the issue and offered a brand-new family of polling techniques called Neighbourhood Expectation Polling (NEP). Each sampled person in NEP answers with details on the percentage of her social network neighbours that have label 1. We took into account the scenarios in which either: 1) The pollster is unfamiliar with the social graph but is nevertheless able to conduct random walks on the graph 2) There are evenly sampled nodes from the unidentified social graph. For cases 1 and 2, two NEP techniques that take use of the relationship paradox kind of system bias were presented.

Theorems 3 through 8 were developed to describe the bias, variance, and mean-squared error of the estimate and how they depend on the underlying network's characteristics (correlation between node labels and degree, expansion, average, minimum and maximum degree, etc.). These findings let a pollster pick the optimum algorithm (in terms of statistical efficiency) and ensure its effectiveness by using previous information about the underlying network. Numerous empirical and simulation data are offered to demonstrate how well the suggested strategies perform when diverse network features are taken into account.

These support the theoretical study and offer perceptions on how the suggested algorithms might function in various scenarios. Both theoretical and experimental findings show that the NEP methods based on the friendships dilemma are able to estimate data with a lesser mean-square error using just a smaller (relative to other approaches) number of respondents.

**REFERENCES:**

[1] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welpe, “Predicting elections with twitter: What 140 characters reveal about political sentiment.” ICWSM, vol. 10, no. 1, pp. 178–185, 2010.

[2] K. J. Gile, “Improved inference for respondent-driven sampling data with application to HIV prevalence estimation,” Journal of the American Statistical Association, vol. 106, no. 493, pp. 135–146, 2011.

[3] A. Dasgupta, R. Kumar, and D. Sivakumar, “Social sampling,” in Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2012, pp. 235–243.

[4] D. Kempe, J. Kleinberg, and E. Tardos, “Maximizing the spread ´ of influence through a social network,” in Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2003, pp. 137–146

[5] S. L. Feld, “Why your friends have more friends than you do,” American Journal of Sociology, vol. 96, no. 6, pp. 1464–1477, 1991.

[6] D. M. Rothschild and J. Wolfers, “Forecasting elections: Voter intentions versus expectations,” 2011.

[7] A. Graefe, “Accuracy gains of adding vote expectation surveys to a combined forecast of us presidential election outcomes,” Research & Politics, vol. 2, no. 1, p. 2053168015570416, 2015.

[8] A. E. Murr, “The wisdom of crowds: Applying Condorcet’s jury theorem to forecasting us presidential elections,” International Journal of Forecasting, vol. 31, no. 3, pp. 916–929, 2015.

[9] A. Graefe, “Accuracy of vote expectation surveys in forecasting elections,” Public Opinion Quarterly, vol. 78, no. S1, pp. 204–232, 2014.

[10] A. E. Murr, ““Wisdom of crowds”? a decentralised election forecasting model that uses citizens local expectations,” Electoral Studies, vol. 30, no. 4, pp. 771–783, 2011.

[11] C. F. Manski, “Measuring expectations,” Econometrica, vol. 72, no. 5, pp. 1329–1376, 2004.

[12] V. Krishnamurthy, Partially Observed Markov Decision Processes. Cambridge University Press, 2016.

[13] V. Krishnamurthy and W. Hoiles, “Online reputation and polling systems: Data incest, social learning, and revealed preferences,” IEEE Transactions on Computational Social Systems, vol. 1, no. 3, pp. 164–179, 2014.

[14] B. Nettasinghe and V. Krishnamurthy, “Maximum likelihood estimation of power-law degree distributions using friendship paradox based sampling,” arXiv preprint arXiv:1908.00310, 2019.

[15] Y.-H. Eom and H.-H. Jo, “Tail-scope: Using friends to estimate heavy tails of degree distributions in large-scale complex networks,” Scientific reports, vol. 5, p. 09752, 2015.

[16] M. Garcia-Herranz, E. Moro, M. Cebrian, N. A. Christakis, and J. H. Fowler, “Using friends as sensors to detect global-scale contagious outbreaks,” PloS one, vol. 9, no. 4, p. e92413, 2014.

[17] N. A. Christakis and J. H. Fowler, “Social network sensors for early detection of contagious outbreaks,” PloS one, vol. 5, no. 9, p. e12948, 2010.

[18] N. Alipourfard, B. Nettasinghe, A. Abeliuk, V. Krishnamurthy, and K. Lerman, “Friendship paradox biases perceptions in directed networks,” arXiv preprint arXiv:1905.05286, 2019.

[19] M. O. Jackson, “The friendship paradox and systematic biases in perceptions and social norms,” Journal of Political Economy, vol. 127, no. 2, pp. 777–818, 2019.

[20] K. Lerman, X. Yan, and X.-Z. Wu, “The “majority illusion” in social networks,” PloS one, vol. 11, no. 2, p. e0147617, 2016.

[21] B. Nettasinghe, V. Krishnamurthy, and K. Lerman, “Diffusion in social networks: Effects of monophilic contagion, friendship paradox and reactive networks,” IEEE Transactions on Network Science and Engineering, 2019.

[22] V. Krishnamurthy and B. Nettasinghe, “Information diffusion in social networks: friendship paradox based models and statistical inference,” in Modeling, Stochastic Control, Optimization, and Applications, ser. The IMA Volumes in Mathematics and its Applications, 2019, vol. 164, pp. 369–406.

[23] E. Lee, S. Lee, Y.-H. Eom, P. Holme, and H.-H. Jo, “Impact of perception models on friendship paradox and opinion formation,” Physical Review E, vol. 99, no. 5, p. 052302, 2019.

[24] J. P. Bagrow, C. M. Danforth, and L. Mitchell, “Which friends are more popular than you?: Contact strength and the friendship paradox in social networks,” in Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017. ACM, 2017, pp. 103–108.

[25] A. Chin, D. Eckles, and J. Ugander, “Evaluating stochastic seeding strategies in networks,” arXiv preprint arXiv:1809.09561, 2018.

[26] V. Kumar, D. Krackhardt, and S. Feld, “Network interventions based on inversity: Leveraging the friendship paradox in unknown network structures,” Yale University, Tech. Rep., 2018.