Social Network Analysis

CS/SDP 361/352
Project Proposal
Hammad Malik hm08298
Mehlab Kashani mk07950

Analyzing Candidate–Party Affiliations in Pakistan: A Social Network Analysis of National Assembly Elections (2013–2024)

Name	Program	year	Email	Role/Responsibility
Mehlab Kashani	SDP	2026	mk07950@st.habib.edu.pk	Analysis and Writing
Hammad Malik	CS	2026	hm08298@st.habib.edu.ok	Data Processing and Visualizations

Project Overview:

Political parties and candidates form the backbone of any democracy. Their dynamics, including competition, collaboration, and shifts in alliance, not only shape election outcomes but also the nature of governance. In Pakistan, these relationships are unconventionally fluid with party switching, new alliances, and independents, which all influence the stability of the political structure. Understanding these movements as to why some candidates remain loyal while others shift sides enables us to evaluate the evolution of Pakistan's democratic institutions. This study employs social network analysis, where candidates and parties are treated as nodes in the network, and edges connect candidates to their affiliated parties in each election

Focusing on three electoral cycles: 2013, 2018, and 2024, the project aims to examine the patterns of loyalty and defection. It aims to construct a bipartite network of candidate–party relationships to analyse the political affiliations. It further aims to map connections that highlight recurring loyalties and the role of independents; apply centrality measures (degree, betweenness, eigenvector, closeness) to identify key parties and candidates acting as bridges across blocs; and detect communities and coalitions using clustering algorithms to identify the existence of any regional or ideological alignments. Lastly, to analyze temporal changes in network structure to understand whether the party system is moving towards consolidation or further fragmentation.

Specific Research Questions:

1. How stable are candidate–party affiliations across Pakistan's National Assembly elections (2013, 2018, and 2024), and what do these patterns reveal about party loyalty, political stability, and external political influences?
2. Which political parties attract the highest number of switching candidates, and how does this reflect their centrality and influence within Pakistan's political network?
3. How has the overall structure of the candidate–party network evolved over the past decade, and what does this indicate about the institutionalization or fragmentation of Pakistan's party system, in light of regional, ideological, and external political shifts?

Data Description

This project draws on electoral records from Pakistan's National Assembly elections, sourced from the Gallup Pakistan Elections Database (1970–2024), which contains 24,585 entries across eleven elections. For our analysis, we focus on the 2013, 2018, and 2024 electoral cycles, comprising roughly 7,000 candidate–party records.

The dataset is in Excel format and captures the relationships between candidates and political parties, along with categorical details such as constituency identifiers (NA numbers), provincial designations, and party classifications. Each record includes variables such as candidate name, party affiliation, constituency, province, vote count, and election year, making it well-suited for bipartite network modelling. In the constructed network, nodes represent candidates (identified by unique name–constituency pairs) and political parties, while edges are electoral affiliations, forming two distinct node sets, one for candidates and one for parties that together define the **network's bipartite structure.** The edge construction will be discussed in more detail further in the proposal.

All data used is publicly available and contains no sensitive personal information. The analysis is conducted with political neutrality, focusing purely on structural and temporal patterns in Pakistan's party system.

Literature review:

The party-candidate affiliation network has a significant impact on political dynamics. Meneghetti et al. (2023) note how party-switching behaviour is prevalent throughout democracies like the US, Australia, Brazil, and many more. To understand these patterns, social network analysis (SNA) is being increasingly employed as a method. It helps to trace the connections between candidates and parties, identify clusters of loyal versus switching members, and reveal which parties attract the most candidates over time. This review examines studies utilising a network-based approach to understand electoral politics, particularly candidate-party affiliation. It also contextualises it within Pakistan's electoral politics, emphasising the need for

social network analysis. By reviewing these articles, it highlights their methodological strengths, limitations and their relevance to the proposed project: *Network Analysis of Candidate-Party Affiliations in National Assembly Elections (2013-2024).*

Bayer & Malang (2025) analyse the impact of party switching on parliamentary behaviour, focusing on Zambia's 11th National Assembly. The study uses interview data of 103 members of the Zambian National Assembly conducted during the 2016 election cycle. The 103 MPs were represented as nodes, with directed ties (edges) indicating political exchange occurrence. The study further uses Exponential Random Graph Models (ERGM) to analyse the network and applies Markov Chain Monte Carlo Maximum Likelihood Estimation (MCMC MLE) to estimate the influence of variables like party affiliation and previous party affiliation on MPs' likelihood to exchange information. It demonstrates how party-switching influences parliamentary interactions by showing how MPs align their social network ties to their current party after switching, as well as maintain and expand connections from their former party.

Similarly, another study by Faustino et al. (2019) estimates the ideology of political parties through party-switching data in Brazil. Using a longitudinal dataset of over two million candidates in Brazil from 1998 to 2018, the authors construct affiliation-switch networks where political parties are represented as nodes, and the edges are formed based on the number of politicians who switched parties between election cycles. The weight of each edge is proportional to the number of switching politicians, showing the intensity of the affiliation change. They applied community detection algorithms, particularly Stochastic Block Model (SBM), to reveal how political parties are grouped based on their ideological similarities and how these groups evolved after politicians switched parties. Further, they also used a beta distribution to model the ideological affinity between parties during their switches. The study revealed how party-switching behaviours have a significant impact on ideological changes within political parties, and parties with similar ideologies cluster together, highlighting that switching follows ideological rather than opportunistic purposes.

Both studies highlight the importance of party-switching in shaping political structure. Bayer & Malang (2025) highlight how it determines political interactions, while Faustino et al. (2019) focus on its impact on ideological alignment. Both studies are directly relevant as they employ Social Network Analysis (SNA) as a methodological approach. They show the relevance of using social network analysis in the proposed project of the candidate-party affiliation network in Pakistan.

While the above studies use social network analysis on topics that are directly relevant to the proposed project, Brito et al. (2020) employ complex network theory to analyse the dynamics of political interactions in the Brazilian Chamber of Deputies. In this network, deputies are represented as nodes and edges are formed based on the similarity in their voting patterns. They

use community detection algorithms, including the Leiden algorithm, to identify political coalitions and analyse the evolution of voting patterns across legislative periods. This study highlights the importance of using social network analysis in understanding the political dynamics.

In the context of Pakistan, candidate party networks are directly relevant considering its political structure. Zhirnov & Mufti (2019) use logistic regression models to examine the likelihood of candidates switching parties. Using datasets of elections from 1993 and 2008, they reveal party-switching patterns, highlighting how switching rates have increased after the 2000s. They further show how party-voter linkages make candidate-party connections more stable. Candidates with personal votes find it easier to switch parties.

In a similar way, Wu & Ali (2020) discuss the transformation of Pakistan's political system from two to a three-party system with the emergence of Pakistan Tehreek-e-Insaf (PTI). They emphasise the rise of PTI and examine PTI's performance from 1997 to 2018, showing how it gained more influence. As a result, it impacted the electoral dynamics, party politics, and government formation. This led to a structural shift in Pakistan's party politics, making it important to study the election cycle of 2013 and 2018 to understand how this transformation impacted the candidate-party affiliation network.

Both studies highlight the importance of studying candidate-party affiliation networks in the context of Pakistan's political structure. Along with this, there is limited research in Pakistan on this topic and particularly using the social network analysis method. Thus, it is crucial to study this network and fill the gap in the literature.

Conclusively, the review highlights how party-switching networks are important and have a significant impact on ideology shifts and political interactions. Although there are election networks based on voter behaviour and prediction that employ social network analysis, there is limited research specifically on party-switching networks, and a significant gap exists in the context of Pakistan. Pakistan is an interesting case study considering its political structure, as a result makes it important to study this network in the context of Pakistan.

Proposed Methodology

This project follows a structured methodology to analyze candidate-party affiliations in Pakistan's National Assembly elections using Social Network Analysis (SNA) techniques. The process is systematically divided into data preparation, network construction, analysis, and visualization phases.

5.1 Data Collection and Filtering

The primary dataset sourced from the Gallup Pakistan Elections Database contains records from eleven National Assembly elections (1970-2024). For focused temporal analysis, we will extract data from three recent electoral cycles:

- **2013 Elections**: Post-democratic transition period featuring established party competition
- **2018 Elections**: Critical power transfer marking Pakistan's second consecutive democratic transition
- **2024 Elections**: Most recent electoral data reflecting current political alignments

This filtering process reduces the dataset from 24,585 total records to approximately 7,000 relevant entries, ensuring computational efficiency while maintaining analytical depth.

5.2 Data Preprocessing

The preprocessing phase ensures data quality and consistency for network construction:

- **Duplicate Removal**: Identifying and eliminating duplicate candidate entries within the same election cycle while preserving multiple entries across different elections to track party switching
- **Name Standardization**: Implementing string matching algorithms to resolve variations in candidate names (e.g., "Muhammad" vs "Mohammad") and party abbreviations (e.g., "PML-N" vs "PMLN")
- **Missing Data Treatment**: Removing incomplete records lacking essential fields (candidate name, party affiliation, or constituency identifier)
- **Edge List Preparation**: Transforming the cleaned data into a bipartite edge list format with source nodes (candidates), target nodes (parties), and edge weights (frequency of affiliation). Single edge per candidate-party pair with weight equal to number of elections together - Party switching creates separate edges to different parties, each with their own weight

5.3 Network Construction

The network will be modelled as a weighted bipartite graph consisting of two distinct node sets:

Node Definition:

- **Candidate Nodes**: Each unique candidate-constituency combination forms a node, accounting for candidates who contest from different constituencies across elections
- **Party Nodes**: Each political party or independent group designation forms a node, including major national parties, regional parties, religious alliances, and independent candidates

Edge Construction:

- **Single edge per candidate-party pair**: Each unique candidate-party relationship is represented by one edge
- **Edge weight calculation**: Weight = number of electoral cycles in which the candidate contested under that party (1 for single election, 2 if in two elections, 3 if all three)

- **No duplicate edges**: If a candidate contests under Party A in 2013 and 2018, this creates one edge with weight = 2, not two separate edges
- **Party switching creates new edges**: If a candidate switches from Party A to Party B, they will have edges to both parties with appropriate weights

Temporal Network Layers:

- **2013 Network**: Baseline network capturing post-transition political structure
- **2018 Network**: Mid-period network showing evolution of affiliations
- **2024 Network**: Current network reflecting contemporary alignments
- **Aggregate Network**: Combined three-period network with weighted edges showing persistence of affiliations

5.4 Network Analysis Techniques

5.4.1 Structural Analysis

Global Network Metrics:

- **Network Density:** Calculates the ratio of actual to possible edges to assess overall connectivity between candidates and parties.
- **Average Path Length:** Measures the average number of steps required to connect any two nodes in the network, indicating how closely linked candidates and parties are within the political structure.
- **Network Diameter:** Determines the longest shortest path between any two nodes, revealing the extent of political fragmentation or cohesion in the network.
- **Degree Distribution:** Examines whether party sizes follow power-law distributions, identifying whether a few dominant parties attract most candidates.
- **Clustering Coefficient:** Captures the tendency of nodes to form tightly knit groups, showing whether candidates affiliated with the same or related parties form local clusters of political association.
- **Bipartite Clustering Coefficient:** Specifically measures the likelihood that two candidates share the same party affiliation, helping identify overlapping political loyalties and coalition tendencies.

5.4.2 Centrality Measures

Party-Level Analysis:

- **Degree Centrality**: Identifying parties with the highest number of affiliated candidates across all elections
- **Weighted Degree Centrality**: Measuring party strength by summing edge weights, capturing both candidate count and loyalty
- **Eigenvector Centrality**: Finding influential parties connected to prominent candidates or other important parties

- **Betweenness Centrality**: Detecting parties that bridge different political segments, potentially indicating coalition builders
- **Closeness Centrality**: Assessing how central a party is to the overall political network

Candidate-Level Analysis:

- **Degree Centrality:** Detects candidates with multiple party affiliations across elections, indicating political mobility or opportunism.
- **Betweenness Centrality:** Identifies candidates who serve as bridges between different political movements or factions, facilitating information or alliance flow.
- **Closeness Centrality:** Evaluates how quickly a candidate can connect to all parties in the network, representing their embeddedness and reach across the political spectrum.
- **Temporal Degree Change:** Tracks how a candidate's party connections evolve over time, capturing shifts in political alignment or strategic repositioning.

5.4.3 Community Detection

- **Bipartite Modularity Optimization**: Applying the Louvain algorithm adapted for bipartite networks to detect communities
- **Political Clustering**: Identifying groups of candidates and parties that form cohesive political blocs
- **Regional Patterns**: Detecting geographic clustering based on constituency-party relationships

5.5 Visualization and Interpretation

Gephi Implementation:

Visualizations will be generated using tools Gephi to depict the bipartite network.

Team Responsibilities and Work Distribution

Task	Responsible Member	Deadline
Data collection and cleaning (filtering 2013–2024 records, formatting Excel sheets)	Hammad Malik (CS)	Nov 2, 2025

Network construction in R (building bipartite edge list and node list)	Hammad Malik (CS)	Nov 2, 2025
Computation of network measures (density, clustering coefficient, average path length, diameter)	Hammad Malik (CS)	Nov 2, 2025
Computation of centrality measures (degree, betweenness, closeness, eigenvector)	Hammad Malik (CS)	Nov 2, 2025
Visualization and temporal analysis in Gephi (2013–2018–2024 layers)	Hammad Malik (CS)	Nov 2, 2025
Interpretation of results (identifying influential parties, switchers, and community structures)	Mehlab Kashani (SDP)	Nov 16, 2025
Writing analytical and discussion sections (linking findings to political context)	Mehlab Kashani (SDP)	Nov 16, 2025
Draft paper preparation and review (M2 Submission)	Both Members	Nov 16, 2025
Final paper and poster design (FP Submission)	Both Members	Nov 30, 2025

References

Bayer, A., & Malang, T. (2025). Late to the party: Do party switchers behave differently in parliament from party loyalists*? International Political Science Review.* https://doi.org/10.1177/01925121251343841

Brito, A. C. M., Silva, F. N., & Amancio, D. R. (2020). A complex network approach to political analysis: Application to the Brazilian Chamber of Deputies. *PLoS ONE, 15*(3), e0229928. https://doi.org/10.1371/journal.pone.0229928

Faustino, J., Barbosa, H., Ribeiro, E., & Menezes, R. (2019). A data-driven network approach for characterization of political parties' ideology dynamics. *Applied Network Science, 4*, 48. https://doi.org/10.1007/s41109-019-0161-0

Meneghetti, N., Pacini, F., Biondi Dal Monte, F., Cracchiolo, M., Rossi, E., Mazzoni, A., & Micera, S. (2023). Predicting party switching through machine learning and open data. *iScience, 6*(10), 107098. https://doi.org/10.1016/j.isci.2023.107098

Wu, X., & Ali, S. (2020). The novel changes in Pakistan's party politics: Analysis of causes and impacts. *Chinese Political Science Review, 5*, 513–533. https://doi.org/10.1007/s41111-020-00156-z

Zhirnov, A., & Mufti, M. (2019). Electoral constraints on inter-party mobility of candidates. Comparative Politics*, 51*(4), 519–537. https://www.jstor.org/stable/26663946