CHAPTER 2
BACKGROUND

Communication means sending data from one to another node within a network. It may involve several other intermediate nodes, which collects data and transfers it to other connected node. The information communicates in a store and forward manner. Network coding has advantages over the traditional store and forward technique of data communication. It increases throughput and capacity of a network. Ahlswede et al. in 2000 proposed network coding technique. This technique studied comprehensively for solving various computer networking problems. Network coding provides improved solutions to several problems of computer networks and its diverse applications. In this chapter we give a brief description of network coding and its applications. Further, we highlight the problems in parallel networks and possible proposed solutions to these problems. This chapter discusses various fields and problems where network coding principle applies.

2.1 Network Coding

Routing problems always open possibilities to improve communication complexity and throughput. The traditional technique used to communicate in a network is by store and forward approach. Research in this field shows that network coding offers better manner to communicate same information in comparison with traditional approach in a network. Network coding allows the intermediate node to encode the received information which increases the maximum data transfer rate. In conventional store and forward technique, only the source node is responsible to encode the data.

All intermediate nodes send the information to other connected nodes. This advantage of network coding leads to a revolution in the research of information theory and network communication. Let us understand the advantage of network coding over store and forward method with the help of an example.
Transmission 1: Node A sends its data to intermediate node.

Transmission 2: Node B sends its data to intermediate node.

Transmission 3: Intermediate node sends information of Node A to Node B.

Transmission 4: Intermediate node sends information of Node B to Node A.

Figure 2.1: Transmission using Store and Forward technique.

In figure 1, two nodes (A and B) communicate with each other using store and forward method. Node A sends data to the intermediate node and then node B sends data to same intermediate node. This data collection at the intermediate node consumes two transmissions.
Further, the intermediate node sends data of A to node B which will take one transmission. Similarly, this intermediate node will transfer data of B to node A in another transmission. This shows that using store and forward method, the data exchanges between node A and B in four transmissions. While using network coding, the data exchanges in three transmissions between node A and B. Figure 2 shows the advantage of network coding over store and forward technique as it reduces the number of transmissions from four to three.

Network coding also provides advantage to increase the capacity of the network. Using the common example of butterfly network (figure 3) we have shown this advantage. Node P_1 and P_2 are the sender node which send [1; 0] and [0; 1] to node P_5 and P_6. While in this communication, node P_3 receives data from both node P_1 and P_2. Now, network coding encode (XORs) this
received information and communicates this information $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ to $P_4$. Thus in this example, network coding reduces the transmissions and increases the capacity of the network. Further this chapter explains the encoding and decoding operations in network coding techniques.

Performance of network is directly based on the flow of data with respect to algorithm. A node without input is called source node and a nodes with one or more inputs is called non-source node. The set of non-source nodes contains the destination node. Now by coding nodes it receives multiple data at a unit of time and results in combined encoded value from these nodes. This encoded value is combined in one data unit and the receiving nodes decode the information to receive the data. Thus, network coding has four major advantages: on the throughput, complexity, robustness and security. Network coding provides efficient methods for several applications and remedy for many problems:

2.1.1 Throughput

Throughput of a communication network is successful delivery rate of a message over a network. One of the major benefits of network coding approach is throughput [2, 3]. Ahlswede et al. showed that network coding provides maximum achievable flow in a network. Noguchi et al. presented the benefits of network coding for high and effective
network utilization. Thus, throughput is the most important utility of network coding. This approach provides increase in throughput by reducing the number of transmissions [1, 2]. Figure 2.1 and 2.2 shows an example which compares network coding and store and forward technique.

2.1.2 Complexity
In several applications network coding has proved that it provides more optimal solution to complex problems [3, 5]. Besides, network coding improves performance by removing or reducing the practical limitations associated with suboptimal solutions [4]. This approach reduces the communication and computation complexities.

2.1.3 Robustness
Network coding is an optimal solution for robustness in several applications which inherits the encoding scheme for data communication in network nodes [6-9]. Network coding provides robustness against the packet loss and link failure in the network. Several solutions are in literature which provide robustness for packet loss in a network, even network coding is most significant solution to this problem. Similarly for link failure (which may be because of several reasons) network coding is efficient mechanism.

2.1.4 Security
The concepts of information theory, like entropy, depend on the communication methods used in a network. Security issues are based on strength of the adversary during and after communication. These problems are addressed by deploying network coding in communication. As security is one of the advantages which network coding provides, it still lacks in some security issues.

The principle of network coding is implemented with network routers and switches [3, 4]. This advantage of network coding is useful with various applications. This chapter discusses the applications and advantages of network coding:
2.2 Applications

The network coding applications are almost in diverse and several fields. The foremost applications are as following:

1: Distributed Storage System [5-8]
2: Content Distribution [9-14]
3: Layered Multicast [15-20]
4: Throughput Enhancement [21-35]
5: Flooding (Broadcast Storm Problem) [36-45]
6: Network Error Correction Code [46-57]
7: Erasure Correction Code [58-74]
8: Loss Tomography [75-83]
9: Topology Inference [84-86]
10: Pollution Attacks [59, 87-94]
11: Eavesdropping [59, 95-99]

2.2.1 Distributed Storage System

The operation performed dynamically need data storage during communication. However, it is difficult to have everything on one disk as disk failure results in losing entire information. So, Dimakis et al. provided a solution to this problem by spreading data on several nodes with redundancy [5]. Several applications follow the distribution storage system. [5] describes an approach that can repair failures in encoded systems by deciding the information to repair this failure. Dimakis et al. have achieved reduced repair bandwidth at the cost of high storage. Wu et al. further characterized tradeoffs between storage and repair bandwidth using a cut based approach [6]. They also presented the approach an algebraic path-weaving technique to prove the existence of codes which achieves the optimal tradeoff. Wu presented in [7] the techniques for constructing network codes. Use of these codes achieves optimal tradeoff between storage efficiency and repair network bandwidth. The size of these codes depends only on the number of nodes at an instant of communication, whereas it is independent of number of failures or repairs. Dimakis et al. have proposed a survey on network codes for distributed storage [8]. In this survey, they presented a review based on erasure coding approach in distributed storage systems to solve the problem of reducing
repair traffic. They analyzed the previous results based on various models of erasure coding. These researches evolved network coding as an effective solution for high storage and failure scenarios.

2.2.2 Content Distribution

Content distribution over network is a huge problem of reliability, security, robustness and resource optimization in several applications. Cohen presented BitTorrent file distribution system which search for Pareto efficiency using tit-for-tat method [9]. Zhu et al. have proposed multicasting in application-layer overlay networks using network coding [10]. This research considerably improves end-to-end throughput of multicast by considering advantages of the overlay network nodes (which are capable beyond basic operations of store and forward method) and the overlay topology. They used network coding in the overlay network nodes to encode and decode data at the message level. However, advantage of the overlay topology to make a 2-redundant multicast graph. Further, Gkantsidis et al. proposed that content distribution of large files over a network (see figure 2.4) is efficient by using network coding principles [11]. Large files distributed using this approach in a dynamic environment. This distribution based on local information without any centralized knowledge of network topology. This makes scheduling large-scale content propagation easier. The network coding approach is also useful as heterogeneous networks do not have any knowledge about the arrival and departure of network nodes. Further, network coding restricts the attack of malicious nodes for content distribution [100, 101].

Chiu et al. studied the use of network coding in P2P (Peer-to-Peer) networks [12]. Based on a simple star network, the maximum achievable throughput is analyzed comparatively between network coded and traditional approach of routing in P2P networks. This work states that, there is no coding advantage using a simple model for P2P content distribution networks. Further Zhang et al. have proposed an approach for peers which are seeking maximization of individual payoff [13]. This distinguishes a coding based P2P content distribution. This work characterizes the traditional P2P approach as a special case of network coding. Further, it is shown that market flexibility to impatient agents is improved using network coding. Ma et al. proposed another network coding approach for P2P content distribution [14]. This approach verifies the dependencies in blocks before transmitting. They also implemented two
other systems: encode and LRF (Local-Rarest-First) and experimentally proved the significance of the approach over traditional approaches.

![Diagram of Content Distribution using Network Coding.](image)

**Figure 2.4:** Content Distribution using Network Coding.

### 2.2.3 Layered Multicast

Multicast is the method to communicate data from one source to multiple destinations. While based on the receivers’ capacity, the layering of each receiver enables flexibility in the process. Network coding provides better throughput for layered multicast and increases the possibility of achieving throughput based on the progression of large size. Cui et al. solved the problem of dynamic media distribution problem using a P2P streaming solution [15]. Zhao et al. have proposed a solution to improve the throughput of an overlay multicast session [16]. Using cache-and-relay and layer-encoded streaming techniques the solution of asynchrony of user requests and heterogeneity of peer network bandwidth are resolved in this work. Chenguang et al. proposed multirate streaming for multimedia data in directed networks using network coding [17]. [18] provide solution for computing the optimal size of each receiver layer. Further in [19] Dumitrescu et al. proposed layered multicast with inter-layer network coding. This proposed work utilizes the full potential of network coding by
multicasting in different data layers. Thus the throughput increases due to flexibility in optimizing the data flow. Applications like video streaming requires multicast at different rates to different receivers [20]. To enable multi-resolution multicast Kim et al. proposed two-stage message passing algorithm which generates network codes for single-source multicast. This work focuses on maximizing the receiving of total layers by all receivers.

### 2.2.4 Throughput Enhancement

Considering figure 2.1 and 2.2, using network coding, in three transmission steps the data exchanges between node A and B. Therefore, in multi-hop network, this approach increases the network throughput and improves the possibility of energy consumption. It is necessary to use the capacity of a wired or wireless network to implement enhancement approaches with maximum opportunities [21]. Ni et al. studied these opportunities of network coding in wireless mesh network [22]. They consider that using network coding what is performance gain while routing in wireless mesh network. COPE analyzes the Routing with Opportunistically Coded Exchanges (ROCX) approach resulting in reduced number of routing transmissions. Further for theoretical perspective, Sengupta et al. analyzed COPE-type network coding in wireless networks for throughput improvements [23]. Chachulski et al. utilizes the advantages of opportunistic routing to further improve network throughput [24]. They proposed MORE (MAC-independent opportunistic routing protocol) approach. MORE enables random mixing of packets before transmitting to ensure the routers to transfer different packets. This approach increases the throughput and robustness in transmissions. While Chaporkar et al. suggested adaptive network coding and scheduling for maximizing throughput in wireless networks [25]. Sagduyu et al. studied network coding advantages over wired networks. The proposed model develops network codes by working with MAC schedulers [26]. [27-35] further studied techniques to improve the network throughput and provide better routing methods.

### 2.2.5 Flooding: Broadcast Storm Problem

Flooding is an advantage to broadcast by which the opportunity to network receiving nodes increases. But the basic broadcasting with flooding creates problem of redundancy, collision etc. and increases the broadcasting cost. This problem is known as Broadcast Storm Problem [36]. Ni et al. identified this problem and analyzed and simulated the results. Further Ni et al. proposed several methods to overcome the problem of redundancy in broadcasting and timing
to broadcast. Peng et al. studied this broadcast storm problem. They proposed an approach to avoid redundant broadcasts by utilizing the topological and statistical information [37]. Several approaches addressed this problem and have proposed various approaches to resolve it. Network coding enables the flooding advantages for its maximum utilization [38-45]. Network coding-based broadcast [102-104] is the solution to the problem of several data packets with several nodes i.e., flooding. Implementation of network coding with flooding reduces the traffic size and opportunities of receiving nodes are still many. Network coding in flooding is also a solution to the broadcast storm problem [105-108].

2.2.6 Network Error and Erasure Correction Code

The traditional error correction codes are not efficient for transmission redundancy. Cai and Yeung have proposed Network Error Correction (NEC) [47, 48] to recover the last packets using network coding. NEC recovers the loss of data packets and the number of links. Besides, Ho et al. have proposed two recovery schemes. First one is receiver based recovery scheme and second is the network wide recovery scheme. Network coding has the utility of spatial redundancy which is used by network erasure correction code while recovering the last packets. Yang et al. proposed an algorithm which constructs network codes that achieves Singleton bound [51]. Further, Matsumoto et al. also proposed an algorithm for constructing linear network error-correcting codes. They defined the association of robust network coding with the network error-correcting codes [52]. Zhang defined the concept of the minimum distance of a network error-correction code. Using two proposed decoding algorithms the performance on MDS code is analyzed [53]. Bao et al. presented Adaptive Network Coded Cooperation (ANCC) for wireless relay networks [54]. Further Koetter et al. have presented the approach for error in random networks and proposed the method for error control in random network coding [55]. Supplementary to this approach, Silva et al. proposed an approach to control error in random network coding. They have constructed the codes using Rank-Metric approach. The tools developed for this approach is directly applicable to random network coding [56]. Guo et al. developed Non-Binary Joint Network-Channel Decoding (NB-JNCD) approach [57] as a reliable scheme for wireless communication. They compared this approach with other traditional approaches for comparative study. As an erasure correction code Koetter et al. have proposed Algebraic Approach to Network Coding [58]. They examined the issues of network capacity in an algebraic framework. Several other
 CHAPTER 2

erasure correction codes mechanisms are proposed with the aim to reduce error in network and increase network capacity [59-74].

2.2.7 Loss Tomography

Network based tomography is the characteristic of the network to send and receive data packets from network edges. Similarly, loss tomography is the characteristic of rate of link loss in the network. Network coding provides advantage in estimating the rate of link loss in the network, identifying these links, tradeoff between bandwidth efficiency and estimated accuracy. The use of distributed network codes infers the link failures location and loss in a network [75]. Using network coding network monitoring for such conditions and loss tomography are further studied for active probing method, coding schemes and loss rate [76-78]. Factor graph approaches are used to visualize the factorization and understand large number of different algorithms. This work developed the understanding of several algorithms by factorizing in graphical form [79]. Mao et al. used this approach of factor graph to visualize the problem of network monitoring for link failures and losses [80]. They considered this approach in wireless sensor networks for monitoring the link loss monitoring. Lin and Gui et al. infer the network monitoring issues using network coding [82, 83].

2.2.8 Topology Inference

In networks with multiple sources and receivers, the communication is iterative and from the network edges. Network coding with topology inference utilizes the network bandwidth and gains the network throughput. Network coding enables reduced transmissions with robust communication with topology inference. Thus this advantage of network coding provides robustness in network communication. Fragouli et al. presented this approach in [84]. Further this approach combines tomographic techniques with network coding [85]. The network coded topological information exactly distinguishes between 2-by-2 subnetwork components. Then the topology is generated and this information is used to merge subnetworks. Yao et al. studied network tomography using network coding for failures. They classified the result as topology estimation and failure localization [86].

2.2.9 Pollution Attack

If in a network, some of the routing nodes are malicious, then these nodes can communicate false information of any combination and further combine with other non-malicious network
nodes and soon this network will become polluted. Such attacks need prevention at early stage. If polluted information is injected in the network then at next hop this information is detected and prevented from creating pollution. Jaggi et al. firstly introduced approach of distributed polynomial-time rate-optimal network codes which works with Byzantine nodes. The algorithms presented targets the various attacking capability of the enemies [87, 94]. In [88], using random network coding in multicast networks Byzantine detection is performed. Gkantsidis et al. proposed the security issues for file distribution in a network using network coding [89]. Several diverse solutions are proposed for the pollution attack on network coding. Another significant approach is signature-based scheme to provide secure network coding. [90-92] have utilized the significance of signature based scheme to resolve the problems of pollution attack. Kehdi et al. proposed Null Keys (a security algorithm) which utilized properties of random linear network coding to detect malicious attacks [93].

2.2.10 Eavesdropping

In a network coded network the packets are communicated from the source node to the receiver. Meantime, if eavesdropping of additional packets occurs, then these packets are incorporated in the network by network coding itself. So, a secure transmission and coding scheme is required to act against this problem. Cai et al. proposed a new model which integrates the concepts of network coding and information security [95]. Further, Feldman et al. showed that making a linear network code secure is equivalent to find a linear code with certain generalized distance properties [96]. Strongly secure linear network coding was proposed by Harada et al [97]. Bhattad and Silva et al. proposed secure linear network coding. Silva et al. have considered the weak security requirements of Bhattad et al. They propose an approach which is independent of network code [98, 99].

Lucani et al. [109] has extended the principle of random linear network coding in consideration of time duplexing channel. They also proposed a method to minimize completion time by finding the packets to be coded in a network. By this research the search time of the coded packets are reduced to an optimal performance. This research ensures the network performance estimating the completion time for packet erasure probability and the packets to be coded. Several problems in the field of network coding are analyzed and resolved by translating in graph theory problem. In [110] the author has converted the problem of linear network coding to a graph theory based problem. In this research Chou et
al. models linear code by taking help of hypergraphs. Authors of this research proposed an algorithm for iterative refinement and in polynomial time this algorithm satisfies linear code constraints. This was the primary systematic approach which solves several network coding problems. Authors have examined several algorithms by converting the problem into graph theory problem and as a result the network bandwidth and computation time is minimized or saved.

Network coding is an advantage for broadcasting in terms of energy efficiency. Network coding improves a factor and this theoretical gain improves the performance. In [41] distributed algorithms are proposed for wireless ad-hoc network scenario. Using simulations it was justified that these proposed distributed algorithms are performing optimally. The proposed work [41] signifies that when network coding is used with wireless ad-hoc networks than the benefits will further increases. Thus network coding enables energy efficient benefits when implemented with wireless ad-hoc network environment. In [40] two algorithms are proposed: a simple XOR based coding algorithm, by which without waiting for more and more coded packets it performs decoding process which makes it NP hard; and a Reed-Solomon based coding algorithm, which is optimal coding algorithm with a limitation that a node will wait until receiving the exact number of coded packets. The first algorithm gives a again upto 45% in comparison to a non coding algorithm and the second algorithm gains upto 61% based on the simulation results. This research proposes a broadcast protocol for mobile ad-hoc networks with network coding i.e. CODEB (Coding-based Broadcast protocol for ad-hoc networks). CODEB consists of three techniques: opportunistic listening; forwarder selection and pruning; and opportunistic coding.

Another approach which considered efficient broadcasting in mobile ad-hoc networks is [45]. This approach considered efficiency problems in mobile ad-hoc networks and provide a solution using network coding and directional antennas. In this approach an additional reduction in the energy consumption while performing broadcast using network coding is achieved using directional antennas. So, this approach combines the advantages of both directional antennas and network coding to achieve efficient broadcast mechanism. An algorithm, efficient broadcast using network coding and directional antennas, was proposed in this research. This algorithm studies the performance variation in a static forwarding node selection. Comparative to traditional broadcast scheme, i.e. CDS-based broadcast, the proposed approach [45] have better performance.
Wireless mesh network is a prominent network in research due to several advantages provided by this network. Broadcasting in wireless mesh network is also an important issue to consider. In [44] using R-code, which is a broadcast protocol based on network coding some one-to-all broadcast scenarios are considered. The backbone of this approach is the minimum spanning tree by which the minimum broadcast overhead and delays. With the help of intra flow network coding an efficient protocol: R-code is presented in this research which reduces the number of transmissions and delay as 14% and 50%.

Other than broadcasting network coding is utilitarian for several security applications in communication. [63] is a prominent application example scenario in which network coding is used for protecting many-to-one wireless flows. Prior to network coding approach, several other approaches like (1+1) protection scheme, (1:N) protection schemes are used for protecting the survivability of many-to-one flows in wireless sensor networks or wireless mesh networks. The major drawbacks with these traditional approaches are the way these approach consume the network resources. [63] proposed a network coding based protection of many-to-one flow in wireless flows which gives advantages over the traditional approaches. In this approach some necessary and sufficient conditions are studied based on network coding. This approach affect the network performance and using the greedy algorithm the scheduling of the transmissions from the source node are proposed. In this approach a polynomial time algorithm is proposed which perform network coding using \{0,1\} coefficients.

Random linear coding is a variant of network coding and several applications are proposed which utilizes the advantages of this approach. In [70] random linear coding is utilized for unicast applications in disruption tolerant network. Application with the factor of buffer space required random linear coding approach to achieve high probability and minimized block delay. This approach is significantly efficient when the constraints like, appropriate token limit choice under bandwidth constraint and under nodal buffer constraint.

Network coding applicability is dependent on the advantages provided by it. Network monitoring for failure occurrence and security is another important issue. [75] utilizes these advantages of network coding for monitoring networks during multicast networks. This approach shows that network coding provides robustness due to the robust distributed
network codes which are utilized to hold information for situation like link failure and losses in the network. This application is an optimal example of network coding applications for failure situations in a network.

Furthermore, network coding applications in the field of security became more prominent. Network coding applicability for content distribution is described in section 2.2.1. [89] states that network coding application for the security issues in file distribution are effectively and efficiently resolved using this approach. [89] proposed and approach with efficient content distribution and robust protection against malicious blocks. The protection schemes in this approach are efficient based on the amount of effort from the attackers. Network coding is providing better throughput for content distribution. [58] is an extension in the previously proposed approach in which multicast networks are studied in terms of network capacity. In [58] the link failure problem is illustrated for network recovery on networks with delay and delay free networks. That means robustness of networking is ensures by this approach.

These application shows that network coding is one of the innovative and vast field with enormous research possibilities. In this thesis we have proposed an application of network coding in the field of parallel communication. Furthermore, this thesis examines applicability of network coding in the parallel communication applications. The subsequent chapters reveal the application of this principle with in the parallel network and are examined based on certain parameters.