

# Abstract

In many real world situations, especially involving multiagent systems and artificial intelligence, participating agents often need to agree upon a common alternative even if they have differing preferences over the available alternatives. Voting is one of the tools of choice in these situations. Common and classic applications of voting in modern applications include collaborative filtering and recommender systems, metasearch engines, coordination and planning among multiple automated agents etc. Agents in these applications usually have computational power at their disposal. This makes the study of computational aspects of voting crucial. This thesis is devoted to a study of computational complexity of several fundamental algorithmic and complexity-theoretic problems arising in the context of voting theory.

The typical setting for our work is an “election”; an election consists of a set of voters or agents, a set of alternatives, and a voting rule. The vote of any agent can be thought of as a ranking (more precisely, a complete order) of the set of alternatives. A voting profile comprises a collection of votes of all the agents. Finally, a voting rule is a mapping that takes as input a voting profile and outputs an alternative, which is called the “winner” or “outcome” of the election. Our contributions in this thesis can be categorized into three parts and are described below.

**Part I: Preference Elicitation.** In the first part of the thesis, we study the problem of eliciting the preferences of a set of voters by asking a small number of comparison queries (such as who a voter prefers between two given alternatives) for various interesting domains of preferences.

We commence with considering the domain of single peaked preferences on trees in Chapter 3. This domain is a significant generalization of the classical well studied domain of single peaked preferences. The domain of single peaked preferences and its generalizations are

hugely popular among political and social scientists. We show tight dependencies between query complexity of preference elicitation and various parameters of the single peaked tree, for example, number of leaves, diameter, path width, maximum degree of a node etc.

We next consider preference elicitation for the domain of single crossing preference profiles in Chapter 4. This domain has also been studied extensively by political scientists, social choice theorists, and computer scientists. We establish that the query complexity of preference elicitation in this domain crucially depends on how the votes are accessed and on whether or not any single crossing ordering is a priori known.

**Part II: Winner Determination.** In the second part of the thesis, we undertake a study of the computational complexity of several important problems related to determining winner of an election.

We begin with a study of the following problem: Given an election, predict the winner of the election under some fixed voting rule by sampling as few votes as possible. We establish optimal or almost optimal bounds on the number of votes that one needs to sample for many commonly used voting rules when the margin of victory is at least  $\epsilon n$  ( $n$  is the number of voters and  $\epsilon$  is a parameter). We next study efficient sampling based algorithms for estimating the margin of victory of a given election for many common voting rules. The margin of victory of an election is a useful measure that captures the robustness of an election outcome. The above two works are presented in Chapter 5.

In Chapter 6, we design an optimal algorithm for determining the plurality winner of an election when the votes are arriving one-by-one in a streaming fashion. This resolves an intriguing question on finding heavy hitters in a stream of items, that has remained open for more than 35 years in the data stream literature. We also provide near optimal algorithms for determining the winner of a stream of votes for other popular voting rules, for example, veto, Borda, maximin etc.

Voters' preferences are often partial orders instead of complete orders. This is known as the incomplete information setting in computational social choice theory. In an incomplete information setting, an extension of the winner determination problem which has been studied extensively is the problem of determining possible winners. We study the kernelization complexity (under the complexity-theoretic framework of parameterized complexity) of the possible winner problem in Chapter 7. We show that there do not exist kernels of size that is polynomial in the number of alternatives for this problem for commonly used voting rules

under a plausible complexity theoretic assumption. However, we also show that the problem of coalitional manipulation which is an important special case of the possible winner problem admits a kernel whose size is polynomially bounded in the number of alternatives for common voting rules.

**Part III: Election Control.** In the final part of the thesis, we study the computational complexity of various interesting aspects of strategic behavior in voting.

First, we consider the impact of partial information in the context of strategic manipulation in Chapter 8. We show that lack of complete information makes the computational problem of manipulation intractable for many commonly used voting rules.

In Chapter 9, we initiate the study of the computational problem of detecting possible instances of election manipulation. We show that detecting manipulation may be computationally easy under certain scenarios even when manipulation is intractable.

The computational problem of bribery is an extensively studied problem in computational social choice theory. We study computational complexity of bribery when the briber is “frugal” in nature. We show for many common voting rules that the bribery problem remains intractable even when the briber’s behavior is restricted to be frugal, thereby strengthening the intractability results from the literature. This forms the subject of Chapter 10.