[[1]](#footnote-2)

Better Allocation to Reduce Voting Queue Length

Enas Mohamed, Hesham Naiem, and Mostafa Izz,

(em\_cmp\_eng , hesham.naiem)@yahoo.com),mostafa.3ez@gmail.com,

Faculty of Engineering, Cairo University

*Abstract*— Providing equitable voting experiences across voting precincts has been noted as an important goal in elections. We seek to provide equity to all voters so that no one particular group of voters is disadvantaged or disenfranchised. This paper uses the average absolute differences of waiting times across all precincts as a performance metric for equity. A simulation-based greedy improvement algorithm is proposed to generate machine allocations. We examine our allocation method using a factorial experimental design, and we conclude that our heuristic outperforms the utilization-equalization method which was used by at least one county in the 2008 presidential election.

*Index Terms*— Voting Queues, Simulation

# INTRODUCTION

Election administration is a service provided to the voting public. As with other service systems, waiting lines are non-value added activities and waiting time is viewed negatively. As discussed in Mebane (2005), waiting times to vote are directly correlated to voter turnout. As waiting times increase, more voters are forced to leave without having cast their ballots due to impatience and other time commitments. Thus, it is very important to design voting systems that result in voters waiting the least amount of time possible. Ideally, we would provide a sufficient quantity of voting machines such that voters would never have to wait to cast ballots. However, due to the expense of new voting machines such as direct-recording electronic

(DRE) machines (also known as “touch-screen” machines), election boards are limited in their ability to procure additional equipment.

Perhaps even more importantly, voting systems should provide equity, which means that we should not design systems that favor some voting groups (defined by geography, voting preference, etc.) by having shorter lines in some precincts than others. Such inequities have been a concern in recent elections. The Department of Justice investigated claims that the Board of Elections in Franklin County, Ohio, “systematically assigned fewer voting machines in polling places serving predominantly black communities as compared to its assignment of machines in predominantly white communities” during the 2004 election (Tanner 2005). While the Department of Justice ultimately found no evidence of systematic decisions to create voting inequities, their report does point to many factors, including specific measures of voter wait times and differences in voting patterns across precincts, which are typically not taken into account under current voting-machine allocation policies. Furthermore, Walter Mebane, Political Science Professor at the University of Michigan, has produced several studies criticizing the Department of Justice’s response to the issues in Franklin County. Professor Mebane concludes from his research, “The allocation of voting machines in Franklin County was clearly biased against voters in precincts with high proportions of African Americans” (Mebane 2006).

Determining an optimal voting-machine allocation is challenging and difficult to solve for several reasons. (1) Voters arrive randomly and according to non-stationary processes to polling locations. There are typically surges in voter arrivals during the morning, noon and evening times due to work schedules (Edelstein 2006). Moreover it is difficult to estimate voter turnout rates prior to Election Day because it depends on many uncontrollable variables (e.g., weather, composition of the voting ballot, etc.). (2) Voter queues may not reach steady state. Ohio law requires that the polls be open 13 hours, plus however much time is needed to accommodate voters waiting to vote at 7:30 pm. Given the limited amount of time that voting precincts are open, the voting queues may still be in a transient state. Such non-stationary arrivals and non-steady-state queues violate the fundamental assumptions of traditional queueing theory. (3) Actual voting scenarios involve considerable computational complexity. There are thousands of polling stations across Ohio and the input variables are stochastic. The result is a large-scale non-linear stochastic optimization problem. Thus, both building model and developing solution methods are challenging endeavors.

We model the voting process using a simulation model that allows us to employ non-stationary arrivals and non-steady state queues. We allocate voting machines to precincts using a greedy improvement heuristic. The objective in our machine allocation is to provide voter equity across precincts. The rest of the paper is arranged as follows. Section 2 provides a review of related literature. Section 3 introduces a performance metric for voter equity and discusses several analysis options for this problem. Section 4 describes the setup of the simulation model and the greedy improvement algorithm (GIA) implemented on the model. Section 5 systematically studies the performance of the simulation-based GIA through an experimental design. Section 6 presents conclusions and future work.

# Motivation

A referendum was held in Egypt on March 19 following the 2011 Egyptian Revolution. For most Egyptians, this was the first genuinely free vote in their lives, so there was a large turnout for the referendum which reached 41.2 % of the 45 million eligible voters. As a result, the queues outside polling stations have continued to grow as Egyptians go out to cast their votes and voters waited patiently for hours in lines. This historical step toward democracy needs to be studied for improvement in the future.

# Problem Definition

Election administration is a service to the voting public. And the waiting lines are a non-value-added activities, and waiting time to vote are directly correlated to voter turnout. Long waiting times usually affect negatively on voter turnout, especially in young countries in democracy like Egypt. This administration process have many challenges like limited number of machines used in voting, the distribution of these machines among different counties and precincts, the variation of turnout for different locations and cultures, different arrival rates during day hours, and other challenges. A simulation based Greedy Improvement Algorithm (GIA) is used to address the problem of voting machines allocation in different precincts.

# Related Work

There are relatively few papers that apply operations-research models to the voting-machine-allocation problem. However, this problem is closely related to queuing-model applications and resource-allocation problems that are common in operations research. Here we will mention only those papers that are most closely related to our application of voting- machine allocation or that reference solution method we use directly.

The only papers of which we are aware that apply operations research to the voting-machine-allocation problem are Edelstein (2006), Allen and Bernshteyn (2006a) and Allen and Bernshteyn (2006b). Edelstein (2006) and Allen and Bernshteyn (2006b) use simulation for their models. Using simulation allows these authors to consider some of the realistic complications in their models including voting-machine failures and non-stationary voter arrivals. However, neither paper explicitly considers voter equity in terms of maintaining equivalent waiting times across precincts. Allen and Bernshteyn (2006a) suggest using queuing models to measure voter waiting times for given machine-allocation policies and to improve allocation decisions. They use simple analytical queuing models to predict average waiting times for voters. Allen and Bernshteyn then suggest an optimization model that uses a minimax objective function to allocate voting machines. Specifically, they suggest allocating machines to minimize the maximum expected voter waiting time across all precincts. The minimax objective is designed to promote voter equity as we discuss above, but there are many other objectives that could be considered. Allen and Bernshteyn (2006a) also do not consider complicating issues such as non-stationary voter arrivals, machine failures, and specific differences in voting-time requirements due to differences in ballot lengths. Furthermore, the authors propose only simple greedy-heuristic solution methods for their models, which can produce significantly suboptimal policies.

There are several simpler methods used to allocate voting machines to precincts that have been used in previous elections. An intuitive and simple method of allocating voting machines used by many election boards is to allocate machines in proportion to the expected number of voters at each precinct (Edelstein 2006). This method ignores any direct models of queuing effects and differences between precincts. At least one county in Ohio used a utilization equalization allocation policy in the 2008 presidential election to allocate voting machines. This method enforces voter equity by equalizing the utilization of voting machines rather than equalizing waiting times of voters. Moreover, the utilization rate is obtained by traditional queuing theory, which assumes stationary arrivals and steady-state operating conditions.

# Analysis Option

## Performance Metric

Simply minimizing total expected waiting time across all precincts is insufficient as this may allow long voter waiting times in some precincts in order to decrease voter waiting time in other precincts. This is undesirable in an election process as we seek to provide equity to all voters so that no one particular group of voters is disadvantaged or disenfranchised.

However, there is no universal way to interpret “equity.” The ideal case is that the expected waiting time in queue at every precinct is the same. But it is generally not feasible to achieve this ideal situation. Therefore, the following metric (the average absolute differences of expected waiting times among precincts) can be used as a proxy for “equity:”

Where N is defined as the total number of voting precincts,, is the number of voting machines allocated to precinct *i*, and is the expected waiting time for voters at precinct *i*.

Thus, the allocations that provide the best “equity” are the global optimal solutions to the following optimization problem:

Where *l* is the set of feasible solutions, and | *l* | is finite.

## Simulation Model vs. Analytical Model

One of the most difficult problems in dealing with voting queues is that voters do not arrive according to a stationary arrival process. Moreover, steady-state may not be achieved in an actual election where voting occurs over a single day and queues begin empty.

Analytical queuing models require strong simplifying assumptions (such as stationary arrivals, steady-state queues, etc.) about the voting system. These models enable us to obtain insights and generate metrics such as expected waiting times very quickly without dedicated simulation software. Moreover, closed-form queuing-model formulas can be used in conjunction with optimization models to determine optimal policy decisions. As an example of the insights offered by such models for this application, an integer-programming-based solution method for this problem using M/M/s closed-form queuing equations shows that voter equity may be compromised if all available voting machines are allocated. The optimal solution to maximize voter equity (as described by (1)) in some scenarios is to not allocate all available voting machines. While this is an interesting, and potentially useful, insight, solving realistically sized problems through an integer program is not generally feasible. Thus, our solution methods described in this paper rely on simulation and heuristic search techniques.

Because of the short time frame of an actual election day, analytical results for the voting-machine-allocation problem require transient queuing analysis with non-stationary arrivals. Obtaining transient information is generally considered much more complicated in comparison to a steady state analysis (e.g., Houdt and Blondia 2005). Roughly speaking, two main approaches have been developed to obtain transient distributions. The first method relies on numerically inverting the Laplace transform or generating function involved (Choudhury, Lucantoni, and Whitt 1994; Hofkens, Spaey, and Blondia 2004; Lucantoni, Choudhury, and Whitt 1994). The second method is based on recursive computations. Others (Ny and Sericola 2003; Lee and Li 1990) combine uniformization techniques to reduce the problem to discrete time and afterward apply a recursive algorithm. Although these methods are effective in obtaining transient distributions related to the initial system behavior, their computational costs grow rapidly when considering later events. Moreover, most of the literature considers only single-server queues. Such limitations of the current analytical results on transient queues weaken the advantages of analytical models, which become more difficult to implement and needs more computational time to obtain results.

Thus, it is natural to turn to stochastic simulation, with its lesser reliance on simplifying assumptions that might render the model questionable in terms of validity. However, we then need to apply proper statistical design and analysis methods in order to deal with uncertainty in the output, and to enable valid and precise conclusions.

# Simulation Analysis

## Description of Basic Polls Queuing Model

### Model Logic

Our simulation model provides the expected waiting time in each precinct for a given number of assigned voting machines. The numbers of DRE voting machines assigned to each precinct are our decision variables.

On Election Day, all polling stations in Ohio open at 6:30 am and close at 7:30 pm. Once a voter arrives to his or her designated precinct, the voter joins a single queue until there is a DRE voting machine available. There are multiple DRE voting machines per precinct. We assume that all DRE voting machines are identical and shared by all voters within a polling station.

### Input Distribution Assumptions and Data Sources

We use a data set based on statistics from the 2004 election in Franklin County, Ohio (available at <http://copperas.com/fcelection/>). Specifically, we fit a normal distribution with mean 1070 and standard deviation 319 to the number of registered voters in each precinct in the 2004 election. For a given number N of precincts (N will be set as a factor in our experimental design below), we generate the number of eligible voters for each precinct independently from this fitted normal distribution. The turnout rate is applied to this, and then the non-stationary arrival pattern is used to distribute these voters’ appearance at the polling station, as described below. We assume that each precinct has the same voter turnout rate.

Ohio Revised Code (O.R.C.) §3501.32 states that on Election Day the polls shall be opened at 6:30 am and shall be closed at 7:30 pm “unless there are voters waiting in line to cast their ballots, in which case the poll shall be kept open until such waiting voters have voted.” Ohio law thus requires that the polls be open 13 hours, plus however much time is needed to accommodate voters waiting to vote at 7:30 pm. Therefore, we allow all queues to clear, but we do not allow any additional voter arrivals after 7:30 pm in our simulations.

On election day, there are “peaks and valleys” of usage by voters depending upon the time of day, the weather, traffic and other variables outside of the control of election staff. Voters do not arrive according to a stationary arrival stochastic process. There are typically surges in voter arrivals during the morning, noon, and evening due to work schedules (Edelstein 2006). Precinct poll workers at the Hamilton County, Ohio, Board of Elections reported an early morning voter rush and lines shorter by day’s end on Election Day 2008. The Voting Experience Survey (Feldman and Belcher 2005; Mebane 2005), which is based on a sample of voters throughout Ohio, provides the percentages of turnout voter arrivals by the time of day (see Table 1). We assume that in each time period the number of arriving voters follows a Poisson distribution. The timing and size of these surges may not be the same across all precincts due to differences in voters’ socioeconomic status, but here we assume that all precincts experience similar arrival patterns.

Table 1: Voter Arrivals by Time of Day

|  |  |
| --- | --- |
| Period of Time | Percentage of Turnout Voters |
| Before 8 a.m. | 20.61 |
| 8 a.m. – 11 a.m. | 27.34 |
| 11 a.m. – 3 p.m. | 24.05 |
| 3 p.m. – 5 p.m. | 13.26 |
| After 5 p.m. | 13.87 |

Ohio law states that voters are allowed up to five minutes to place their vote (Anthony et al. 2004). However, anecdotal evidence suggests that this law is rarely, if ever, enforced. Actual voter service times will depend on the length of the ballot - in particular, the number of issues on the ballot, which generally require the most time for voters to read and on which to make a choice. To determine approximate voter service times we use data from the 2006 Ohio gubernatorial election with six issues on the ballot. We fit a distribution to the data of voting times read from a sample of Election Systems & Software machines in this election and found that a gamma distribution with the scale parameter of 1.05 and the shape parameter of 5.71 fits the data acceptably. We assume that voting times in every precinct follow this same gamma distribution.

For now, we assume that voting machines are perfectly reliable (i.e., there are no voting-machine failures). We also

assume that all available voting machines must be allocated to precincts as this is often the policy used in practice, and will reduce the system-wide voter wait time.

## The Greedy Improvement Algorithm

### Algorithm Description

We use a heuristic solution method to allocate voting machines to precincts. Our method combines a simple greedy heuristic to reduce expected voter wait times with a local improvement search where we use (1) as our objective function. We refer to this solution method as the greedy improvement algorithm (GIA), which consists of two phases: Phase I is a simple greedy heuristic and Phase II is a local improvement search.

### Implementation

It is not difficult to implement the GIA in a standard programming language (e.g., C, VBA, and C++) once the expected wait times in any precinct () are obtained. However, extensive effort is required to build simulation models to obtain the expected wait times using standard programming languages. On the other hand, it is easy to obtain () in simulation modeling software (e.g., Arena), but difficult to implement the GIA, because simulation-modeling software is not designed as a programming language.

We preferred to use a general purpose programming language to implement the simulation model (namely C# .Net) with the help of open source .Net simulation library called

# Experimental Design and Results

In order to examine the performance of the simulation-based RA and compare it with the greedy improvement algorithm (GIA), we set up experimental design using four factors: Voting Time, Number of Precincts, Ratio of the number of machines to the number of precincts, and Allocation Strategy (see Table 2). The result of this experimental design is the “equity” metric given in (1).

Table 2: Factors and Values for Experimental Design

|  |  |
| --- | --- |
| Factors | Possible Values |
| Number of Precincts | 20 – 30 - 50 Precincts |
| Voting Time  (Scale Parameter of Gamma Distribution) | 0.58 - 1.05 |
| #Machines/#Precincts | 2 - 3.6 |
| Allocation Strategy | RA - GIA |

## Factors and Values

To get the voter turnout rate we used the data based on statistics from the 2004 election in Franklin County, Ohio.

We fit a Weibull distribution with Shape Parameter α=6.9514 and Scale Parameter β=60.884 to turnout percentage in each precinct in the 2004 election.

We assume that the voting service time in every precinct follows a gamma distribution with shape parameter of 5.71. One level for this factor is set using data from the 2006 Ohio gubernatorial election by setting the gamma distribution scale parameter to 1.05. Allen and Bernshteyn (2006a) use a mean service time of voting of 3.33 minutes. We use this as the mean for the other level of our Voting Time factor. We do not have data for the voting times with mean of 3.33 minutes, but we assume such voting times also follow a gamma distribution with shape parameter 5.71 and set the distribution’s scale parameter to 0.58 to match this mean voting time.

We set the two values of Number of Precincts to be 20 precincts, 30 precincts, and 50 precincts.

We set the two level of Ratio of the number of machines to the number of precincts to 2 and 3.6.

To obtain the allocations for each combination of design factors, we run the RA and GIA using the React.Net Library. We use 50 replications for each scenario so that the 95% confidence-interval half width will be less than 10% of the average waiting time in a precinct.

There are 12 design points in total (see Table 3). We run 50 replications for each design point. Different random number streams are used for different design points to ensure that the design points are independent.

## Results

The results of the experimental design (see Table 3) show that the GIA statistically significantly outperforms the UEM in 14 out of 16 different treatment combinations of the other four factors, ties in one scenario (Turnout Rate = 0.56, Scale Parameter = 1.05, Size of County = 20 precincts and Ratio of the number of machines to the number of precincts = 3.6), and underperforms UEM in a single scenario (Turnout Rate = 0.72, Scale Parameter = 1.05, Size of County = 20 precincts and Ratio of the number of machines to the number of precincts = 3.6).

The running of simulation for the RA and GIA methods shows that the GIA method takes very long time to run all iterations compared to RA, this is because Phase 1 of the GIA runs the simulation number of times equivalent to the number of DRE machines to be allocated in one replication. For example if we have 100 DRE Machines and 50 replications, the simulation will run for about 100\*50 times in Phase 1 only which is a lot of time. So the RA outperforms the GIA in the speed of simulation.

Figure 1 shows the scenario with Turnout Rate = 0.72, Scale Parameter = 1.05, Size of County = 50 precincts and Ratio of the number of machines to the number of precincts = 3.6. It displays 95% confidence intervals and the range of the response values against the allocations provided by the GIA and the UEM. It clearly shows that the GIA is statistically significantly better in this scenario.

Figure 2 case which is not better.

# Conclusions

The occurrence of long lines in elections depends on many unpredictable factors and is difficult to control. One way to explore this problem is through simulation which can be useful in the election process and can result in good recommendations to the allocation of DRE machines across precincts.

In our simulation model we used non-stationary voter arrivals, transient queues, and different turnout rates across all precincts.

The machine allocation RA method is proposed and illustrated using an example from the 2004 election in Franklin County, Ohio. Then it is compared to the GIA method [ ].

The RA method is shown to offer potential advantages, both through the reduction of equity metric (the average absolute differences of waiting times across all precincts) and it takes smaller time in simulation compared to the GIA.

# [Limitations and] Future Work

An incomplete list of additional topics for future study is as follows:

1. Include more realistic cases such as heterogeneous precincts to the simulation model, which have different voter-arrival patterns, and different distributions of voting times due to ballot length and nature of voter. Also considering the voting-machine failures in the model which occurs in real life.
2. Explore the elections in the developing countries such as Egypt, where the elections don’t have a technological infrastructure for voting because it is not practical or present. It uses manual technique for voting, tallying, or verification of paper-ballots. These factors lead to the lack of necessary data which can be used to simulate elections in developing countries. Studying the elections there, find a way to simulate it, and applying the methods mentioned in the literature could help in strengthening the democratic process in developing countries.
3. Commercial software could be developed based on the RA or any other method on the literature. This software will help governments take decisions about election allocations that result in voters waiting the least amount of time possible and provide savings to the money spent on election process.

Acknowledgment

References

1. Allen, T. T., and M. B. Bernshteyn. 2006a. Mitigating voter waiting times. Chance 19 (4): 25–36.
2. Allen, T. T., and M. B. Bernshteyn. 2006b. optimal voting machine allocation analysis. In Analysis of May 2006 Primary Cuyahoga County, Ohio, ed. S. Hertzberg, 71–89. Election Science Institute. Available via <http://moritzlaw.osu.edu/electionlaw/news/documents/DREAnalysisforMay2006Primary.pdf>
3. Anthony, A. A., M. F. Colley, K. E. Marinello, and C. C. Petree. 2004. Election 2004: A report to the community. Franklin County Board of Elections Report. Available via <http://centralohiovoters.home.att.net/elect2004.htm>
4. Muer Yang, Michael J. Fry, W. David Kelton. ARE ALL VOTING QUEUES CREATED EQUAL? in Proceedings of the 2009 Winter Simulation Conference
5. Edelstein, W. A. 2006. New voting systems for New York - long lines and high cost. Technical report, New Yorkers for Verified Voting Report. Available via [http://www.nyvv.org/ doc/voterlines.pdf](http://www.nyvv.org/%20doc/voterlines.pdf)
6. Feldman, D., and C. Belcher. 2005. Voting experience survey. In Democracy at Risk: The 2004 Election in Ohio. Democratic National Committee. Available via <http://www.democrats.org/pdfs/ohvrireport/section03.pdf>
7. Houdt, B. V., and C. Blondia. 2005. Approximated transient queue length and waiting time distributions via steady state analysis. Stochastic Models 21:725–744.

# Work Division

|  |  |
| --- | --- |
| Team Member | Activites |
| Enas Mohamed,  em\_cmp\_eng@yahoo.com | 1. Search for ideas and conferences 2. Draft of Project Proposal (State of the art, and References) |
| Hesham Naiem Mamoun,  Hesham.naiem@yahoo.com | 1. Search for ideas and conferences 2. Draft of Project Proposal (Target Conference, Motivation, and Potential Contribution) |
| Mostafa Mohamed Izz,  Mostafa.3ez@gmail.com | 1. Search for ideas and conferences 2. Draft of Project Proposal (Problem Statement, and Proposed work and plan) |

1. Manuscript received June XX, 20XX; revised July XX, 20XX. (Write the date on which you submitted your paper for review.) This work was supported in part by the U.S. Depart­ment of Com­merce under Grant BS123456 (sponsor and financial support acknowledgment goes here). Paper titles should be written in uppercase and lowercase letters, not all uppercase. Avoid writing long formulas with subscripts in the title; short formulas that identify the elements are fine (e.g., "Nd–Fe–B"). Do not write “(Invited)” in the title. Full names of authors are preferred in the author field, but are not required. Put a space between authors’ initials. [↑](#footnote-ref-2)