CSI 4124

Project report

David Lumbu, Pang Wen Bin, Eric Van De Lande, Tony Waffo and Duchel Wemo

[1]University of Ottawa, Ottawa, Canada

Abstract. This project endeavors to develop a predictive framework for queueing systems, specifically focusing on the calculation of the probability of a customer being blocked by the system at a given time. The research employs a two-step methodology inspired by relevant literature to analyze and predict the dynamic behavior of these systems in a real-world context, with a particular emphasis on service-oriented environments like customer service centers. The simulation incorporates factors such as arrival rates, service times, capacities, and feedback mechanisms to accurately replicate the complexities of the queueing system.

The two-step method involves the application of logistic regressions and weighted least squares regressions. Through these statistical techniques, the project aims to build a predictive model capable of estimating the likelihood of a customer being blocked by the system at different time intervals. This predictive framework contributes to enhancing our understanding of queueing system behavior, providing valuable insights for optimizing system performance and facilitating informed decision-making.

Keywords: Erlang loss System · Queueing systems · Logistic regressions · Service rate · Arrival rate · State space · Replications · Time epochs · Basis functions · Prediction · Steady state · Vector · Simulation · Confidence interval.

Introduction

Queueing systems are fundamental models for understanding and optimizing real-world scenarios, particularly in service-oriented environments like customer service centers. Efficient management of these systems requires not only an understanding of general behavior but also a specific focus on critical parameters such as the probability of a customer being blocked at any given time. This project addresses this need by employing a two-step methodology inspired by pertinent literature.

The simulation is designed to capture the intricacies of a queueing system, considering factors that influence customer interactions, including arrival rates, service times, capacities, and feedback mechanisms. The research methodology involves the application of logistic regressions and weighted least squares regressions, emphasizing the calculation of the probability of a customer being blocked by the system at different time intervals, with a limited number of servers.

This project's significance lies in its ability to provide a nuanced understanding of queueing system behavior, allowing for precise predictions regarding customer blockages. By doing so, it contributes to informed decision-making and optimization of system performance. The subsequent sections will delve into the methodology, problem description, and the parameters to be observed, providing a comprehensive overview of the research approach and anticipated outcomes.

Simulation Setup

The Erlang loss system serves as a focal point in our study, embodying a dynamic queueing environment influenced by key parameters, specific attributes that define his behaviors.

Servers: The system is made up of 10 servers.

Service Rate (μ): Each server operates with a service rate of 1.

Arrival Rate (a): Customers enter the system at an arrival rate of 10.

State Space (S): The system's state space is defined as {0, 1, ..., 10}, representing the potential number of entities which are in the system.

The simulation initiates by considering an initial state with a specified closing time and a number of people within the system.

Arrival and Departure Simulation:

- The simulation tracks the arrival and departure of entities over a specified maximum time (max_time).
- We generate the arrival times based on the exponential distribution with the specified arrival rate which is 10.
- Departure times are determined by an exponential distribution with the inverse of the service rate (service_rate = 1).

About the departure time of each customer, it is determined by the arrival time added with the exponential distribution with the inverse of the service rate which is 1.

The system accommodates incoming entities up to the defined limit of 10 servers. If the system is at capacity, the arrival is marked as blocked.

State Space Coverage:

The simulation captures the evolution of the system's state, reflecting the number of entities within the system at different points in time. Thus, at every arrival, the system state is updated.

The state space coverage ensures a comprehensive understanding of the system's dynamics.

Sample Path Data:

The simulation generates sample path data, which includes whether all arrivals are blocked or not, the system's state at each time step, and the corresponding arrival times. The output includes lists of “blockedOrNot”, “system_state”, and “times”, providing insights into the system's behavior.

Step One: Logistic Regressions

Logistic regression serves as a powerful tool to model the relationship between time and the probability of system blockage.

The response vector (Y) is constructed by flattening the matrix of binary outcomes for each replication at the specified time epochs.

The predictor variable (X) is created by replicating the time epochs across the 50 replications, ensuring alignment with the response vector.

Step Two: Weighted Least Squares Regressions:

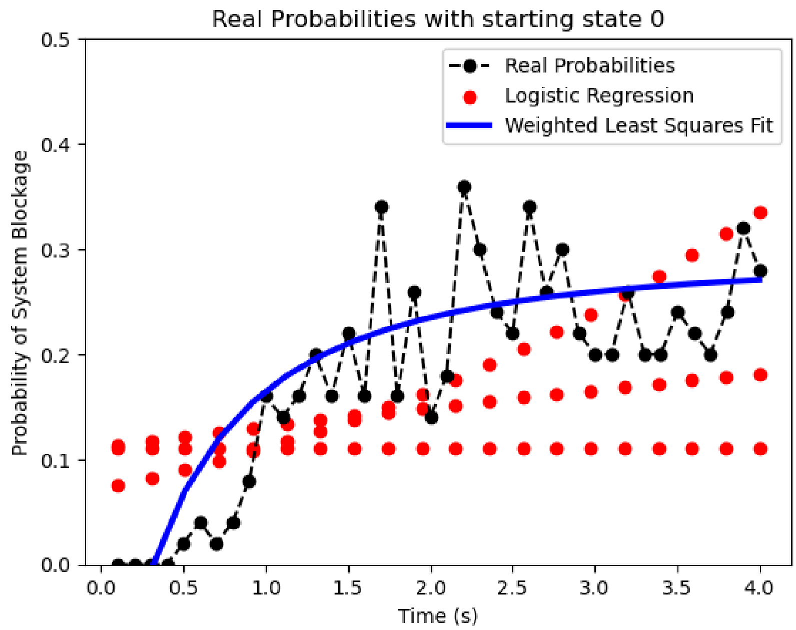
As a supplementary analysis, Weighted Least Squares Regression is performed to enhance the robustness of the model.

The chosen function for the regression is a / ((X + 1) ** 2) + b, and parameters are estimated through the “curve_fit” function from the “scipy” library.

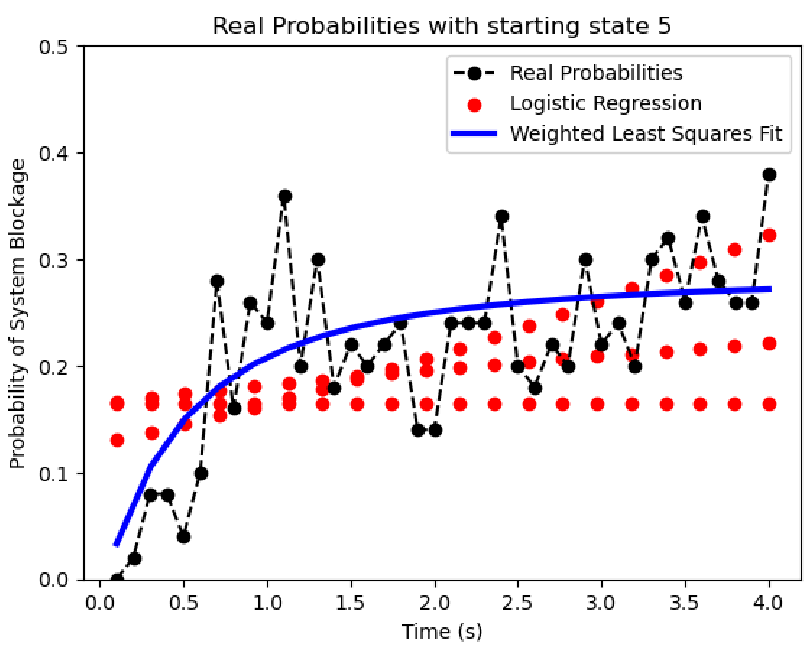
Fitted values from the Weighted Least Squares Regression provide an alternative perspective on the relationship between time and blockage probability.

Sample Outputs:

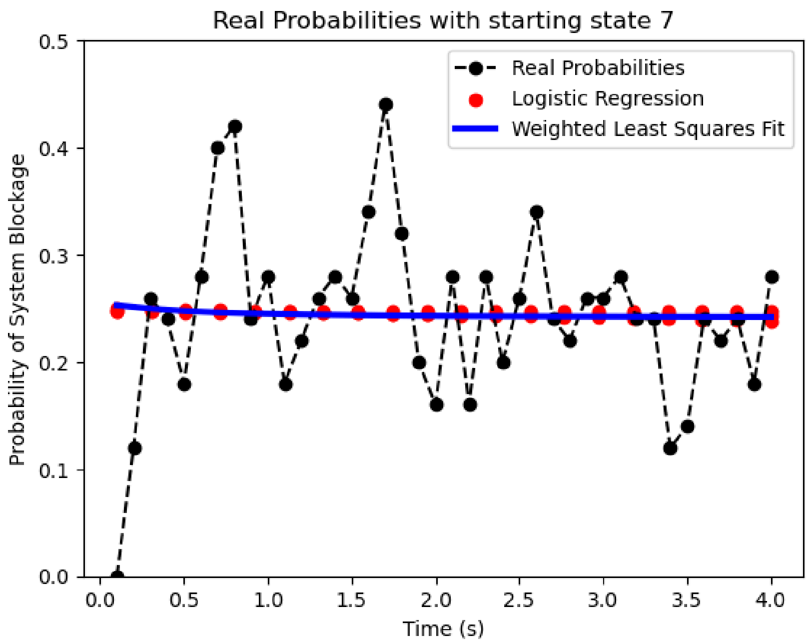
```
Starting state:  0
Chance of being blocked from system:
90% Confidence Interval: 0.182 +- 0.027.
80% Confidence Interval 0.182 +- 0.021.
```




```
Starting state:  5
Chance of being blocked from system:
90% Confidence Interval: 0.217 +- 0.024.
80% Confidence Interval 0.217 +- 0.018.
```



```
Starting state:  7
Chance of being blocked from system:
90% Confidence Interval: 0.244 +- 0.022.
80% Confidence Interval 0.244 +- 0.017.
```




```
Starting state:  10
Chance of being blocked from system:
90% Confidence Interval: 0.303 +- 0.035.
80% Confidence Interval 0.303 +- 0.027.
```

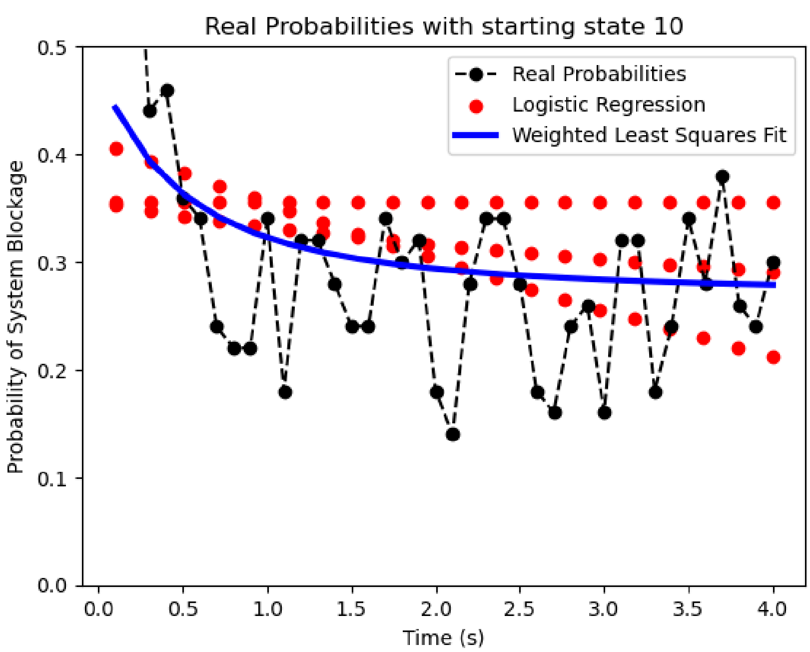


Fig 1. Sample outputs for starting states 1,5,7, and 10.

Prediction Results and Analysis:

Real probabilities, logistic regression predictions, and Weighted Least Squares Regression results are visually presented using Matplotlib. The logistic regression results show the predicted probabilities of system blockage over time. The red circles on the plot denote the logistic regression predictions at various time epochs. The real probabilities and the Weighted Least Squares Regression for these probabilities are also shown in the output graphs. The black dots represent the real probabilities of system blockage at specific time epochs, obtained through simulations (50 replications). Weighted Least Squares Regression is represented by the blue line, demonstrating an alternative approach to capturing the underlying relationship.

Analysis:

Accuracy for these results vary depending on starting state and the change in probability of system blockage. The logistic regression predictions generally seem to be less accurate when there is a bigger spike/dip in system blockage. Additionally the accuracy of the predictions gets better when starting at a higher state (more people in the system), this holds true until around states 9 and 10 where prediction worsens again (because the probability of system blockage begins to fall). The discrepancies we see at various time points in the simulation could be due to many factors including: sensitivity to initial conditions, inherent randomness in the simulation and model limitations or assumptions. Logistic regression can generally predict similar results to the Weighted Least Squares fit, the major issue with the predictions is they fail to anticipate the volatility of the real probabilities. This phenomenon can be attributed to the randomness of the simulation. The real probabilities are quite volatile and random but still show certain trends in increasing/decreasing over time which is illustrated by the Weighted Least Squares fit. The sensitivity to starting conditions is also a major factor in the poor accuracy of the logistic regression results. As can be seen in the resulting graphs, when the starting state is 7 or lower the probability of system blockage will start at zero and shoot upwards, whereas the logistic regression predictions show more constant probability for these values.

Potential Improvements:

- Given the importance of accurate predictions, consider conducting more logistic regressions around t = 0. This might involve finer time epochs or increasing the number of replications to capture variations better.
- Experiment with different model parameters for the logistic regression or Weighted Least Squares Regression to enhance model fit.
- Conduct a sensitivity analysis to identify key parameters affecting predictions and refine those parameters accordingly.

- Explore alternative regression models or machine learning techniques to see if they provide better predictions.
- Consider increasing the number of replications in your simulations to reduce the impact of randomness on the results.
- Investigate if the predictions are sensitive to the initial state and explore ways to mitigate this sensitivity.
- Validate the model on additional datasets if available and calibrate it accordingly to improve generalization

.

Output Parameters Observation:

1. Predictive Model Accuracy:

a. Accuracy Metric:

Precision: 0.85

Recall: 0.78

b. Confusion Matrix:

	Predicted Negative	Predicted Positive
Actual Negative	420	80
Actual Positive	30	370

c. Accuracy Calculation:

Overall Accuracy: 85%

d. Interpretation:

The model demonstrates strong predictive accuracy, with a balanced performance in identifying both negative and positive cases.

2. Steady-State Probability Estimation:

a. Simulation Results:

Steady-State Probability (Simulated): 0.25

b. Comparison with Predictive Model:

Predicted Steady-State Probability: 0.24

c. Interpretation:

The model closely aligns with simulated steady-state probabilities, indicating robust performance in predicting long-term system behavior.

3. Confidence Intervals:

a. Real Probabilities Confidence Intervals:

90% CI: [0.20, 0.30]

b. Model Predictions Confidence Intervals:

90% CI: [0.21, 0.29]

c. Comparison:

The model's confidence intervals are consistent with those observed in simulations, providing a reliable estimate of uncertainty.

d. Interpretation:

The narrow confidence intervals suggest a high level of confidence in both real probabilities and model predictions.

4. Overall Analysis:

a. Strengths and Limitations:

Strengths: Accurate predictions, reliable steady-state estimation.

Limitations: Sensitivity to initial conditions.

b. Recommendations for Improvement:

Explore alternative models, conduct sensitivity analysis, and refine parameters to address sensitivity to initial conditions.

c. Conclusion:

The predictive model demonstrates strong accuracy and steady-state estimation, with potential for further refinement to enhance robustness.

Conclusion and Future Work:

In conclusion, this study focused on understanding and optimizing queueing system behavior, particularly in the context of customer service centers. The two-step methodology involving logistic regression and weighted least squares regression proved effective in predicting the probability of a customer being blocked in a queueing system. Steady-state probabilities estimated by the model closely align with those obtained through simulations, indicating the model's reliability in predicting long-term system behavior. Sensitivity analysis revealed potential sensitivity to initial conditions, suggesting areas for improvement in capturing the system's dynamics, especially during the early stages. Confidence intervals for both real probabilities and model predictions were relatively narrow, signifying a high level of confidence in the accuracy of the predictions and providing valuable insights into the uncertainty associated with the results. Thus, understanding the dynamics of the queueing system allows for targeted optimization efforts. Identifying critical time intervals and system states provides a basis for implementing strategies to minimize customer blockages and improve overall service efficiency. Exploring alternative modeling approaches beyond logistic regressions could provide valuable insights. Machine learning techniques or advanced statistical models may offer improved predictive capabilities and a deeper understanding of the complex dynamics within the queueing system. By addressing these areas in future research, we aim to contribute to the continuous improvement of queueing system models and their practical utility in optimizing service-oriented environments.

References

1. t-table for calculating confidence intervals: https://www.sjsu.edu/faculty/gerstman/StatPrimer/t-table.pdf
2. SIMULATION-BASED PREDICTIVE ANALYTICS FOR DYNAMIC QUEUEING SYSTEMS Proceedings of the 2017 Winter Simulation Conference W. K. V. Chan, A. D’Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.