Call Center Capacity Planning

The main topics of the thesis are theoretical and applied queueing theory within a call center setting. Call centers have in recent years become the main means of communication between customers and companies, and between citizens and public institutions. The extensively computerized infrastructure in modern call centers allows for a high level of customization, but also induces complicated operational processes. The size of the industry together with the complex and labor intensive nature of large call centers motivates the research carried out to understand the underlying processes. The customizable infrastructure allows customers to be divided into classes depending on their requests or their value to the call center operator. The agents working in call centers can in the same way be split into groups based on their skills. The discipline of matching calls from different customer classes to agent groups is known as skills-based routing. It involves designing the routing policies in a way that results in customers receiving a desired service level such as the waiting time they experience. The emphasis of this thesis is on the design of these policies. The first paper, Queues with waiting time dependent service, introduces a novel approach to analyzing queueing systems. This involves using the waiting time of the first customer in line as the primary variable on which the analysis is based. The legacy approach has been to use the number of customers in queue. The new approach facilitates exact analysis of systems where service depends on the waiting time. Two such systems are analyzed, one where a server can adapt its service speed according to the waiting time of the first customer in line. The other deals with a two-server setup where one of the servers is only allowed to take customers who have waited a certain fixed amount of time. The latter case is based on a commonly used rule in call centers to control overflow between agent groups. Realistic call center models require multi-server setups to be analyzed. For this reason, an approximation based on the waiting time of the first in line approach is developed in the paper Waiting time dependent multi-server priority queues, which is able to deal with multi-server setups. It is used to analyze a setup with two customer classes and two agent groups, with overflow between them controlled by a fixed threshold. Waiting time distributions are obtained in order to relate the results to the service levels used in call centers. Furthermore, the generic nature of the approximation is demonstrated by applying it to a system incorporating a dynamic priority scheme. In the last paper Optimization of overflow policies in call centers, overflows between agent groups are further analyzed. The design of the overflow policies is optimized using Markov Decision Processes and a gain with regard to service levels is obtained. Also, the fixed threshold policy is investigated and found to be appropriate when one class is given high priority and when it is desired that calls are answered by the designated agent class and not by other groups through overflow.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Contributors: Nielsen, T. B.
Publication date: Mar 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
(ImM-PHD-2009-223).
Electronic versions:
phd223_tbn.pdf
Source: orbit
Source-ID: 251790
Research output: Research › Ph.D. thesis – Annual report year: 2010