Pergamon

0305-0483(95)00050-X

Omega, Int. J. Mgmt Sci. Vol. 24, No. 2, pp. 195–203, 1996

0305-0483/96 $15.00 + 0.00

A Field Service Support System Using a Queueing Network Model and the Priority MVA Algorithm

HT PAPADOPOULOS

University of the Aegean, Samos, Greece

(Received January 1995; accepted after revision September 1995)

In many companies, the field service (FS) department plays an important role, contributing up to 40% to the company's total revenue. FS managers have to cope with conflicting objectives: (a) to maintain a high level of customer service and (b) to keep the spares inventory level as low as possible. Therefore, they need tools to analyse the impact of their decisions on customer service and inventory cost. Such a tool is presented in this paper. We have developed a closed queueing network model, similar to one due to Waller, incorporating priority classes of customers via the application of the priority mean value analysis (PMVA) algorithm, developed by Shalev-Oren *et al.* This model has been applied to the FS organization of the Greek subsidiary of a multinational computer company, and it has proved to be very efficient from the computational point of view, thus constituting a powerful tool for the FS managers. Copyright © 1996 Elsevier Science Ltd

Key words—queueing, performance evaluation, field service, decision support systems, strategic planning

1. INTRODUCTION

MANY INDUSTRIES, multinational and national companies set as their first priority 'customer satisfaction', as they know that only in this way they can maintain or increase their market share. The sales forces sell products but the battle does not end there. On the contrary! At that point a new battle begins, namely, that of service maintenance and support of those products, which fall under the responsibility of the field service or customer service departments. Currently, large companies offer multi-vendor integrative solutions to their customers, thus undertaking the responsibility of the service maintenance of all those different products which may come from different manufacturers and, therefore, making the management of the field service a more difficult task.

Expenses in the budget of a field service (FS) department are the FS spares inventory and the wages of the FS department employees, plus some general expenses for travelling and training. Some of the important decisions that the FS manager has to make are related to the management of the spare parts inventory, to the number of FS engineers and to the allocation of the FS engineers to the company's customers.

The motivation for this work has been a paper by Waller [3] in conjunction with the author's experience on all the issues relevant to the FS organization, since he has served as the logistics and administration manager and the FS contracts sales manager of a multinational computer subsidiary in Athens, Greece. First, we make some additional remarks on Waller's presentation and secondly we suggest the application of a closed queueing network (CQN) model implementing the priority mean value analysis (PMVA) algorithm developed by Shalev-Oren *et al.* [2]. We have applied both models, the Baskett, Chandy, Muntz and

Palacios (BCMP) (Waller's CQN model), and the PMVA (the proposed model), to the data of the FS organization of the multinational computer subsidiary in Athens, and we have seen that both models perform quite satisfactorily with the latter giving more accurate (for this FS organization) results. Unfortunately, for reasons of confidentiality, numerical data and results cannot be presented.

This paper is organized as follows. In Section 2, we present the structure of a real FS organization, bearing in mind that of the (multinational) computer subsidiary. We also make some additional remarks on Waller's presentation, relevant to the function and operations of this organization. In Section 3, we develop our model. In Section 4, we present our conclusions and recommendations for further research. Finally, in the Appendix, the analysis of the CQN model is given together with the implementation of the PMVA algorithm.

2. THE STRUCTURE OF AN FS ORGANIZATION OF A COMPUTER COMPANY

In this section, we describe briefly the structure and the operation of the FS organization of a computer (multinational) subsidiary, as these have motivated and formed the basis for the development of our FS support system. We believe that the structure and operation of FS departments in other industries are quite similar (e.g. photocopier, communication companies, etc.). For this reason, we strongly recommend the application of the proposed model, described in the next section, to these companies as well, with slight modifications and adjustments. In a computer company (subsidiary of a multinational, or a national of medium size) the FS organization consists mainly of two departments:

- (a) the FS sales department, which brings revenue to the organization by selling service contracts and other services (such as preparation of computing rooms to accommodate the computing systems) to the customers. Another source of revenue is the revenue transfer coming from the sales department as a certain percentage on the sales of hardware (H/W) and software (S/W) products, covering the one-year period during which the products are under guarantee; and

- (b) the technical support department, which consists of the engineers (technicians), for both H/W and S/W products. This department, together with the wages of the administrators, the secretaries, the engineers, and the call handler(s), the various travelling and training expenses, constitute the expensive part of the FS organization.

The operation of the technical support department is as follows. The customer who has a problem calls the call-handler in the FS support department and this person logs the call, checks the contract of the customer and passes the information to the dispatcher who arranges for an engineer to take care of this call, according to the terms and conditions of the contract of this customer.

There are various classes of customers, depending on the type of their contract with the company. A typical classification, for example, is:

(1) *Class-1 customers*: These are the few customers with 24-hour coverage and they are assigned the highest priority. The company is obliged to service these customers immediately by having an engineer on stand-by during non-working hours, even at weekends and on holidays. An example of such a customer is a production plant (e.g. cement producer) which maintains 3 shifts per 24 hours.

(2) *Class-2 customers*: These are high-priority customers too, and also few. Their contract specifies that an FS engineer must go and fix their problem within 2 hours of the call. A representative example of such a class of customers is a bank which wants to maintain a reliable on-line system.

The class 1 and 2 customers and their repair times are small enough that these customers experience no waiting, which is justified through analysis of real data.

(3) *Class-3 customers*: These are also priority customers, whose contracts specify that an FS engineer must go and fix their problem within 4 hours of the call. An insurance company may belong to this class of customers.

(4) *Class-4 customers*: These are the normal customers. Their contracts specify that an FS engineer must go and fix their problem within 8 hours of the call. The majority of the customer population of the company belong to this class (e.g. Universities, research institutes, various private commercial companies).

(5) *Class-5 customers*: These are customers with an *elementary* service maintenance contract, that cannot afford a normal contract and try to be covered somehow at least from the spare parts point of view. According to their contract, an FS engineer must go and fix their problem within 16 hours of the call. Of course, the cost of such a contract is lower than for a normal contract (say 65% of the price of a normal contract). The company tries not to sell such contracts, and therefore only a few customers belong to this class.

(6) *Class-6 customers*: These are customers *without* a contract, the so-called *per-call customers*. The rules of the company dictate the FS engineers should give them very low priority. Depending on work-load, they are obliged to go and fix their problem within a week, over-charging them, of course, in order to force them to sign a normal service maintenance contract.

The FS manager together with the FS unit managers (managers of the H/W and S/W engineers) are obliged to service all these customers, always trying to be consistent with the terms and conditions of the contracts.

One of the problems the managers face is how to allocate the engineers to the various customers. Usually, the allocation is done depending on the customer class, and the type of system (super, mini, PCs—UNIX, DOS, VMS, etc.) and the part of the computer (CPU, peripherals, discs and tapes, printers, terminals). The case where more than one engineer is specialized in a specific area, part of the computer or type of operating system is very common, for back-up reasons and in order to easily dispatch the calls among them.

Usually, FS managers, restricted by a budget, do not have the luxury of hiring new engineers. Instead, they invest more money in the training of existing staff in different areas (e.g. both for mini systems and PCs, and H/W and S/W products) in order to decrease the idle time of the engineers and increase their utilization, efficiency and productivity in general.

Inventory investment and spare parts stock management are important factors affecting the probability that an engineer will not have the necessary parts (this is denoted by p_E by Waller [3]). We have applied Figs 2–5 of Waller to the data of the Greek subsidiary of the computer company and we have confirmed their suitability. Briefly, these figures of Waller depict graphically the effect of p_E on the percentage of customers which are 'up' for various levels of the (response) delivery time; the effect of p_E on the probability that the FS engineer is idle; the maximum number of customers vs p_E with a service level of 95%; and the percentage of operational customers for different levels of the number of customers assigned to each FS engineer.

In this work, it has not been our aim to deal with stock control models in order to see how these affect p_E; this may be the subject of further research.

Another crucial decision the FS manager has to make is if it is worthwhile to *vendorize* various operations and activities of the whole FS organization business, e.g. to subcontract the repair of all the PCs and terminals and printers, avoiding in this way the mass of the daily calls, keeping the annual cost of repair fixed. Of course, in that case, the FS manager must be very careful about the quality of the service offered by the sub-contractor, as the main objective of the company is customer satisfaction. Strict control of the sub-contractor is needed for this to succeed. Again this issue is beyond the scope of our research and is not considered.

3. MODEL DEVELOPMENT AND APPLICATION

The model that best suits the FS organization of the computer company described in the previous section is a closed queueing network (CQN) with priority scheduling, implementing the priority mean value analysis (PMVA) algorithm, developed by Shalev-Oren *et al.* [2]. This CQN is given in Fig. 1, below, and can accept R classes of customers and $K(r) \geqslant 1$ customers of class r, $1 \leqslant r \leqslant R$. In our case, $R = 6$.

The total number of customers, N, is given by

$$N = \sum_{r=1}^{R} K(r), \qquad (1)$$

and there are M nodes in the network with $J(m) \geqslant 1$ parallel identical servers (the FS engineers) at node m, $0 \leqslant m \leqslant M - 1$.

Node 0 is a fictitious queueing station accommodating all the 'up' customers of the company (something like 'negative' customers), consisting of customers of all classes whose system/option is up and running. Notation related to the mean service rates (or mean service times) for all four nodes of the PMVA network is given in the Appendix.

Reading Fig. 1, and along the lines of [3], when a breakdown occurs, the 'up' customer becomes 'down' and enters node 3 or node 1, depending on the class r to which it belongs. If $r = 1$ or 2, service is immediate at node 3, by a return to node 0, as an 'up' customer. If $r = 3, \ldots, 6$, the customer enters node 1, waiting for the FS engineer's first visit. The customer is serviced by one of the $J(1)$ FS engineers at this node, then returns to node 0. However, there is a probability p_E that the FS engineer does not have (from the company's warehouse) those spare parts, necessary to complete the repair. Then the engineer places a P/1 (emergency) order with the materials planning department for the part(s), and leaves to service another customer. The original customer, only for modelling purposes, enters node 2 to wait the delivery of the part. When the P/1 order arrives at the warehouse, the FS engineer is notified by the warehouse attendant. The

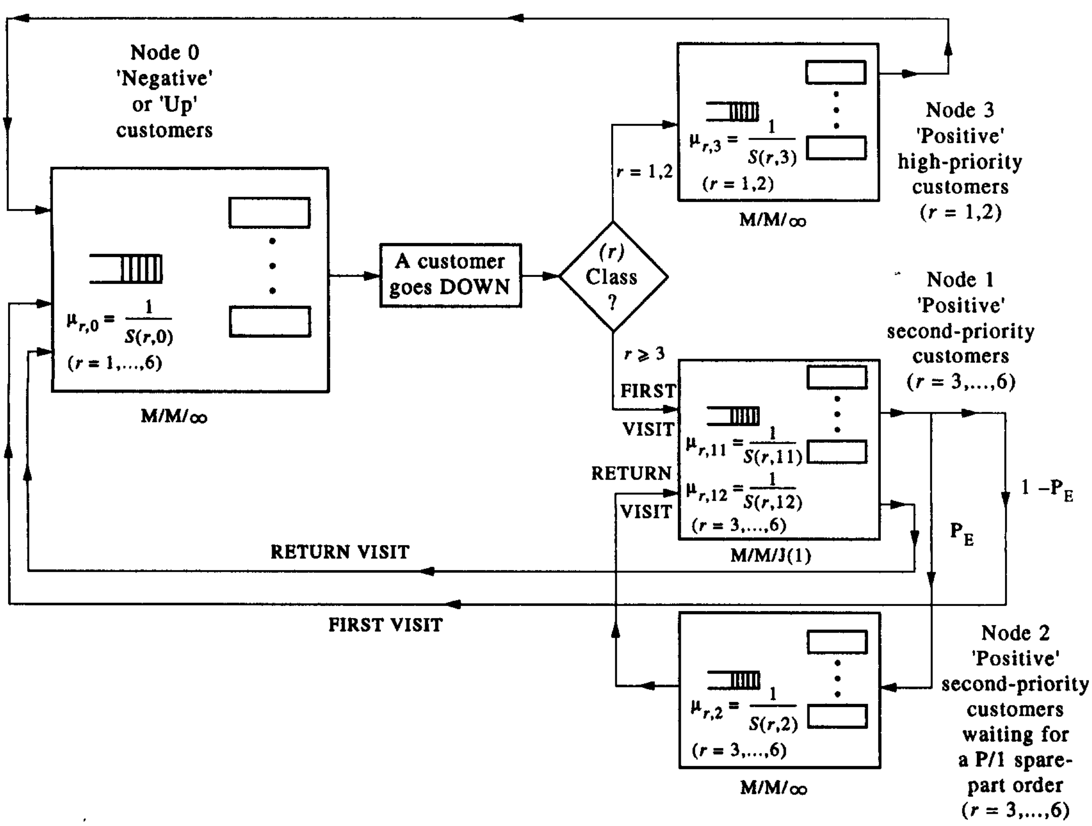

Fig. 1. A PMVA network for the FS organization.

customer re-enters node 1, waiting for the FS engineer's return visit. Upon the arrival of the engineer, the repair is completed and the customer returns to node 0 (as an 'up' customer).

Keeping the notation of Shalev-Oren *et al.* [2], SD(m) denotes the service discipline at station m, with SD = AS (ample server or infinite service or self-service) or FCFS (first-come-first-served) or HOL (head-of-the-line). For each node m having HOL discipline, $PR(r, m)$ denotes the priority assigned to customer class r. The model allows several classes of customers to have the same priority at station m and within each priority, service is FCFS. A certain customer class may have distinct priorities at different stations.

The service (repair) time of a class r customer at a server in station m is assumed to follow an exponential distribution with mean service time $S(r, m)$.

The routing matrix is given for each customer class r and, from this, the equilibrium distribution $\{P(r, m)\}$ is calculated. The $P(r, m)$s are nonnegative and may be scaled arbitrarily (see [2]), being proportional to the mean number of visits to station m by a customer of class r.

The two basic variables that are computed from this model are:

- $G(r, m)$: the mean number of arrivals (or equivalently, departures) per unit time, i.e. the throughput of class r customers at station m;
- $W(r, m)$: the mean time spent by a class r customer waiting in queue when it visits station m.

These variables are calculated from the solution of a $2MR$ non-linear system of equations (see [2]) given in the Appendix.

Then, from the above variables, various interesting performance measures are derived, such as,

- $Q(r, m)[Q(m)]$ = the mean number of class r customers [or all customers (of any class), respectively] on queue at station m,
- $U(r, m)[U(m)]$ = the utilization of a typical server (i.e. FS engineer) at station m, due to class r customers [or due to all classes of customers],
- $T(r, m)[T(m)]$ = the mean sojourn time (queueing plus repair time) of a class r customer [or all customers] at station m,
- $N(r, m)[N(m)]$ = the expected number of class r customers [or all customers] at station m (on queue and in service),
- $G(m)$ = the throughput at station m, and
- $QS(r)[QS]$ = the mean number of class r customers on queue somewhere [or in the total system].

All these variables are defined mathematically in the Appendix, where the proposed model is analysed and further notation is given.

Now, concerning the network topology, we have examined a few alternatives, from which we have decided to adopt only two (after having run the PMVA algorithm and compared the results against the real ones from the data of the FS organization of the Greek subsidiary of the computer company).

The first topology (T1) is the same as Waller's. Briefly describing this topology, which is depicted in Fig. 1 of Waller's paper [3], it models the FS support system as a Baskett, Chandy, Muntz and Palacios (BCMP) queueing network with two classes of customers. One class includes those customers waiting for an initial visit from the FS engineer and the other models those customers whose the spare parts that have been ordered via an emergency (P/1) order, have arrived and they are waiting for the FS engineer to return to install them.

The second topology (T2) is shown in our Fig. 1, and has an additional node 3, which is analysed as an $M/M/\infty$ queueing system (according to Kendall's notation). This node was added to accommodate the class-1 and class-2 customers (e.g. those of the highest priority). These few customers are served immediately. For them, there is dedicated stock (the company keeps kits only for them), and special stock (option-swap material). Moreover, they do not have to wait for an engineer to finish a job somewhere else and then to come. So, there is no waiting time for them in the queue (for either a spare part to come from the supplier or for an engineer to be delayed in arriving). Empirical evidence confirms that modelling node 3 as a self-service queueing system ($M/M/\infty$) is almost always correct. The only case, where these two classes of the highest priority customers are not served immediately, is when the spare part (which is always available at the warehouse or at the customer's site) is dead-on-arrival, i.e. although it is brand-new and sealed, it fails to function correctly. This case is very rare, and the FS manager tries to find a solution for the 'critical' customer.

All the other classes of customers (3–6) are served at node 1 as in [3], but with the HOL service discipline, as these classes of customers are assigned different priorities, according to their contracts. Customers belonging to the same class are serviced according to FCFS service discipline.

We applied the data of the Greek subsidiary of the multinational computer company, using the above two topologies: [(T1): Waller's model with $J(1)$ FS engineers at node 1, instead of 1, and FCFS service discipline, and (T2): the modified Waller's topology with an extra node (3) and HOL service discipline at node 1], and utilized the PMVA algorithm. We found that both alternatives gave numerical results of remarkable accuracy, on comparing them with the real data, concerning the various performance measures (such as the expected number of customers at each node and the expected utilization of the FS engineers, or, equivalently, the expected idle time of the FS engineers). The second topology provided results with slightly better accuracy [2–6% deviation for (T2) and 5–10% deviation for (T1), respectively, from the real data]. However, for class-6 (per-call) and class-5 (with an elementary service contract) customers, the results from the (T2) configuration were worse than the respective results obtained from (T1). The explanation for this is that the errors in estimating $W_0(r, m)$ (see the Appendix) have a greater impact in the case of the HOL service discipline than in the FCFS case. This happens when the data [$K(6)$ and $K(5)$, i.e. the population of class-6 and class-5 customers] are small, which is true for these two classes of customers.

We also ran the BCMP model, proposed by Waller, and we found that the results were also quite satisfactory but worse than ours (with an average absolute difference from real data, 15–20%).

Concerning the travel time of the FS engineers (or equivalently of the customers in our model), we have not applied an extra dummy node (say $M + 1$) in the network, although the PMVA algorithm provides this facility, but we have included it in the service (repair) time of the respective node. We believe that this is reasonable, especially with the use of a dedicated courier to move the spare parts from the warehouse to the customer and back, allowing the engineers to move only from customer to customer, without being obliged to spend time on travelling to the warehouse and back to the customers. Addition of an extra node (the transporter) is more applicable to the modelling of flexible manufacturing systems, where the material handling system plays an important role in the whole operation. However, it may be possible to model travelling time by a separate node of the CQN, when a dedicated courier is not used by the company and the engineers spend much time on travelling, especially in cities with heavy traffic.

4. CONCLUSIONS AND FURTHER RESEARCH

On applying the priority mean value analysis (PMVA) algorithm, introduced by Shalev-Oren *et al.* [2], we have developed a closed queueing network model to form the basis for a field service support system. By running this model with the data of the FS organization of a subsidiary of a multinational computer company in Athens, Greece, we have found that it gives very accurate results. The above model proved more accurate than Waller's [3] model. In addition, it has the advantage that it allows one to estimate more performance measures such as the mean sojourn (repair) time of a customer of any class. Moreover, due to the inherent convergence criterion of the PMVA algorithm, namely the utilization of the FS engineers (see Appendix), the FS manager can control the idle time of the FS engineers, ensuring high utilization of resources. Also, the proposed CQN model is more realistic, since it allows the detailed modelling of customers with different priorities.

Although we have focused on the application of the proposed CQN model to the small Greek subsidiary of the computer company, we strongly believe that this model is applicable to many other companies with similar structure of their field service organizations, such as, among others, the photocopier and the communications industries. It would be interesting to see whether the numerical results obtained from larger companies are equally satisfactory. Possible deviations would have to do with the impact of errors in estimating the mean time spent by a class r customer in the queue when visiting the repair station, especially with the use of the HOL service discipline.

A quite interesting and useful area for further research would be the development of a simple, easy to use (by the FS manager who does not need to know anything about queueing theory), total system cost model, incorporating all the possible decision variables, such as the number of FS engineers required, their allocation to the various customer classes, the effect of the spares stock policy on the level of service, the FS engineers' utilization or the percentage of their idle time, the cost of training them, how much time a customer waits from the time of a breakdown until the problem has been fixed (total sojourn/repair time), and if the terms and conditions of their contract are being met, etc.

APPENDIX

Here, we first give the notation used in the PMVA network and the mathematical definition of the various performance measures that are derived from this model, and then we present the system equations of this model. The details of their derivation are omitted as they may be found in [2].

Notation

$N = \Sigma_{r=1}^{R=6} K(r)$ = the total number of customers (of all 6 classes) in the system;

p_E = the probability that the FS engineer does not have the necessary spare parts to complete the repair during the first visit;

$S(r, 0)$ = the mean time between failures (MTBF) for a customer of class r ($1 \leqslant r \leqslant R = 6$) (at node 0). This is usually given in hours by the manufacturer and it may be derived statistically from real data;

$S(r, 11)$ = the mean service time for a first visit of a customer of class r ($r \geqslant 3$) at node 1, by a FS engineer. This usually includes travelling time, the time for the diagnosis of the problem and the repair time;

$S(r, 12)$ = the mean service time for a return visit of a customer of class r ($r \geqslant 3$) at node 1, by a FS engineer. This usually includes travelling time, the time for the replacement of the spare part that arrived from the P/1 order and any testing of the equipment;

$S(r, 2)$ = the mean delivery time of a P/1 (emergency) order, for a customer of class r ($r \geqslant 3$) at node 2. This time varies, depending on the geographical location of the supplier (for a European supplier = 18–36 h, for a U.S.A supplier = 3–4 days);

$S(r, 3)$ = the mean service time of the high-priority customers (class $r = 1, 2$) at node 3. This time includes no waiting time on queue, but time just for travelling, diagnosis of the problem, repair, and testing.

Definition of the performance measures

For $1 \leqslant r \leqslant 6(=R)$ and $0 \leqslant m \leqslant 3$, it holds:

$$Q(r,m) = G(r,m)W(r,m),$$

$$Q(m) = \sum_{r=1}^{R} Q(r,m),$$

$$U(r,m) = \frac{G(r,m)S(r,m)}{J(m)} (\leqslant 1),$$

$$U(m) = \sum_{r=1}^{R} U(r,m) \ (\leqslant 1),$$

$$T(r,m) = W(r,m) + S(r,m),$$

$$T(m) = \sum_{r=1}^{R} \frac{T(r,m)G(r,m)}{G(m)},$$

$$N(r,m) = G(r,m)T(r,m),$$

$$N(m) = \sum_{r=1}^{R} N(r,m),$$

$$G(m) = \sum_{r=1}^{R} G(r,m),$$

$$QS(r) = \sum_{m=0}^{M-1} Q(r,m),$$

$$QS = \sum_{r=1}^{R} QS(r).$$

The system equations

The network depicted in Fig. 1, is a PMVA network and the system equations for all the service disciplines (SDs) [ample service (AS), FCFS and HOL] are analysed in detail in [2]. Here we give only the equations for SD = AS, FCFS and HOL, as nodes 0, 2 and 3 are modelled as self-service centers, while node 1 is modelled as a $M/M/c/\mathrm{HOL}$ multi-server queueing system with HOL service discipline, and within each class, customers are serviced according to the FCFS discipline.

The $2MR$ equations for the Gs and Ws are the functional mean value analysis (MVA) equations (see [1, pp. 78–85, and [2]), with modifications of the queueing time estimates due to the parallel servers and the HOL service discipline.

In all the formulae below, r takes values from 1 to $6 = R$ and m lies between 0 and $3 = M - 1$.

The first MR equations are for the Gs:

$$G(r,m) = \frac{P(r,m)K(r)}{\sum_{k=1}^{M} P(r,k)[W(r,k) + S(r,k)]}, \quad (2)$$

$$= \frac{K(r)}{W(r,m) + \theta(r,m)}, \quad (3)$$

where

$$\theta(r,m) = S(r,m) + \left\{ \sum_{\substack{k=1 \\ k \neq m}}^{M} \frac{P(r,k)}{P(r,m)} [S(r,k) + W(r,k)] \right\}. \quad (4)$$

We have used expression (3) as the values for $W(r,m)$ are not computed directly from (2), because iterations sometimes lead to utilizations above unity.

These MR equations are the standard *exact* equations based upon Little's formula for the queueing network as a whole.

The remaining MR equations calculate *approximate* values for the queueing time $W(r,m)$. These equations are dependent on the service discipline of the node.

- When the service discipline is AS:

$$W(r,m) = 0. \quad (5)$$

- For the FCFS service discipline, $W(r,m)$ is given by

$$W(r,m) = W_0(r,m) + \left(\frac{1}{J(m)}\right) \times \left(\sum_{k=1}^{R} G(k,m)W(k,m)S(k,m) - G(r,m)W(r,m)S(r,m)/K(r) \right). \quad (6)$$

Equation (6) expresses the mean queueing delay as the time $W_0(r,m)$ needed to clear one customer from service [if all $J(m)$ servers are busy at the time of arrival], plus the expected time to clear the queue seen by the last class r arriving customer at node m. The term multiplied by $1/J(m)$ is the standard MVA queue workload estimate including the $1/K(r)$ correction, since the mean queue length observed by an arriving customer of class r to a node is roughly equal to the time average queue length at the service station with the arriving customer removed from the system. This correction is exact for $K(r) = 1$ and is asymptotically correct for the other limiting case of very large $K(r)$.

For $J(m) = 1$, $W_0(r,m)$ is calculated using the MVA estimate for the residual service time (see source references in [1] or [2]):

$$W_0(r,m) = \sum_{k=1}^{R} [G(k,m)S(k,m)^2] - G(r,m)S(r,m)^2/K(r). \quad (7)$$

When $J(m) > 1$, Shalev-Oren *et al.* used the following expression:

$$W_0(r,m) = b(r,m)DT(r,m), \quad (8)$$

where $b(r,m)$ denotes the probability that an arrival of a class r customer finds all $J(m)$ FS engineers at node m busy, and $DT(r,m)$ is the mean delay time until first departure from service, if an arrival of class r to service station m finds all $J(m)$ servers (FS engineers) busy. For an approximation of these two quantities, see [2].

- When the service discipline is HOL, the mean waiting time in the queue of station m (node 1 for our model), for an r-class customer, is given by

$$W(r,m) = (1/2)\left[C(r,m) + \sqrt{(C(r,m))^2 + 4H(r,m)}\right], \quad (9)$$

where

$$H(r,m) = \frac{W_0(r,m)\theta(r,m)}{D(r,m)[D(r,m) - E(r,m)]}, \quad (10)$$

$$C(r,m) = \frac{\dfrac{W_0(r,m)}{D(r,m)} + \dfrac{[K(r)-1]S(r,m)}{J(m)}}{[D(r,m) - E(r,m)] - \theta(r,m)}, \quad (11)$$

and

$$D(r,m) = 1 - \frac{1}{J(m)}\sum_{k=1}^{R} G(k,m)S(k,m), \quad PR(k,m) < PR(r,m), \quad (12)$$

$$E(r,m) = \frac{1}{J(m)}\sum_{\substack{k=1\\k\neq r}}^{R} G(k,m)S(k,m), \quad PR(k,m) = PR(r,m), \quad (13)$$

with $PR(r, m)$ denoting the priority of an r-customer at node m, and $\theta(r, m)$ given by equation (4). Finally, the $W_0(r, m)$s are computed, using equations (7) and (8) for $J(m) = 1$ and $J(m) > 1$, respectively, in the same way (as for the FCFS service discipline).

Now, having derived the Gs and Ws from the solution of the $2MR$ simultaneous non-linear equations (3) and (5), for SD = AS, equations (3) and (6), for SD = FCFS, and equations (3) and (9), for SD = HOL, we have easily calculated the various other performance measures (given above), which are both station-measures and system-measures, per customer class, $r = 1, \ldots, R(=6)$, and aggregate (all classes of customers), as well, something that makes the proposed model very attractive.

To effect the solution procedure we have followed the instructions of [2] for the solution of the $2MR$ simultaneous non-linear equations. The interesting point is that for the convergence of the algorithm, we have maintained for the appropriate m (our node 1, see Fig. 1):

$$0.85 \leqslant \sum_{r=1}^{R} \frac{G(r,m)S(r,m)}{J(m)} \leqslant 0.95, \quad (14)$$

which is realistic, as this has to do with the utilization of each server (of our node 1). This is equivalent to controlling the FS engineer's idle item to be as little as possible (15–5%), and this can be as tight as the FS manager wants to make it.

Another stopping criterion employed was the following:

$$|W_k(r,m) - W_{k+1}(r,m)| \leqslant \epsilon, \quad (15)$$

where k is the iteration counter and ϵ is a pre-specified threshold (e.g. $\epsilon = 1 \times 10^{-5}$). The maximum number of customers per class r or in total [$K(r)$ or $N = \Sigma_{r=1}^{R} K(r)$, respectively], that may be assigned to a FS engineer, keeping the customer level of service (LOS) say at least 95% (which is the usual LOS used), is (similar to Waller's) the largest $K(r)$ or N, that satisfies the inequality

$$N(r,0) \geqslant \frac{95K(r)}{100}, \quad (16)$$

or

$$N(0) \geqslant \frac{95N}{100}, \quad (17)$$

respectively.

ACKNOWLEDGEMENTS

The author would like to thank the two anonymous referees for their helpful comments that led to this final version of the first draft. Many thanks are also due to Professor George Flessas for a critical reading of the manuscript.

REFERENCES

1. Papadopoulos HT, Heavey C and Browne J (1993) *Queueing Theory in Manufacturing Systems Analysis and Design*. Chapman & Hall, London.
2. Shalev-Oren S, Seidmann A and Schweitzer P (1985) Analysis of flexible manufacturing systems with priority scheduling: PMVA. *Ann. Opns Res.* **3**, 115–139.
3. Waller AAW (1994) A queueing network model for field service support systems. *Omega* **22**, 35–40.

ADDRESS FOR CORRESPONDENCE: *Professor Hrissoleon T Papadopoulos, Department of Mathematics, University of the Aegean, GR-832 00 Karlovassi, Samos, Greece.*