2016

Τμήμα βιομηχανικής διοίκησης και τεχνολογίας

Μεταπτυχιακό πρόγραμμα σπουδών στη διοίκηση Logistics

Εφαρμογή Θεωρίας Ουρών Αναμονής σε Υπηρεσίες:(1) Τηλεφωνικού Κέντρου και (2) Νοσοκομείου

ΕΠΙΒΛΕΠΩΝ: ΠΑΠΑΔΟΠΟΥΛΟΣ ΧΡΥΣΟΛΕΩΝ
ΕΛΙΣΑΒΕΤ ΚΑΛΑΪΤΖΗ

Περιεχόμενα

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στα πλαίσια του Μεταπτυχιακού Προγράμματος «Διοίκηση Logistics» του τμήματος Βιομηχανικής Διοίκησης και Τεχνολογίας του Πανεπιστημίου Πειραιώς.

Σ’ αυτό το σημείο θα ήθελα να εκφράσω τις ευχαριστίες μου στον καθηγητή μου κ. Χρυσολέοντα Παπαδόπουλο , Καθηγητή στις Ποσοτικές Μεθόδους στη Διοίκηση Παραγωγής, Διοίκηση Logistics και Εφοδιαστικών Συστημάτων για την βοήθεια, εμπιστοσύνη και την εκτίμηση που μου έδειξε.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια και τους φίλους μου για όλη την ηθική και ψυχολογική υποστήριξη που μου προσέφεραν καθ όλη τη διάρκεια των σπουδών μου.

Καλαϊτζή Ελισάβετ

Περίληψη

Η αναμονή σε ουρές αναμονής είναι ένα φαινόμενο το οποίο έχουν βιώσει όλοι στη σύγχρονη κοινωνία και είναι ιδιαίτερα δυσάρεστο. Σκοπός λοιπόν αυτής της εργασίας είναι η ανάλυση των ουρών αναμονής στην θεωρία αλλά και στην πράξη με την παρουσίαση δύο παραδειγμάτων σε συστήματα εξυπηρέτησης, ένα σε ένα τηλεφωνικό κέντρο μίας τράπεζας όπως και στα επείγοντα ενός νοσοκομείου.

Στο πρώτο κεφάλαιο παρουσιάζεται τη θεωρία των ουρών αναμονής σε συστήματα αναμονής. Παρατηρούμε κάποια από τα βασικά μοντέλα ουρών αναμονής όπως και κάποιες κατανομές αλλά και θεωρήματα που θα μας φανούν χρήσιμες στην πορεία.

Στο δεύτερο κεφάλαιο αναλύεται η μεταβλητότητα και πως αυτή επηρεάζει τα συστήματα ουρών αναμονής και υπολογίζεται ο μέσος χρόνος αναμονής σε ουρές με έναν αλλά και με περισσότερους εξυπηρετητές.

Στο τρίτο κεφάλαιο παρατηρείται η απώλεια της απόδοσης λόγω της μεταβλητότητας και υπολογίζεται πόσοι πελάτες χάνονται από το σύστημα αλλά και τι ποσοστό του χρόνου το σύστημα βρίσκεται σε κατάσταση πληρότητας.

Στο τέταρτο κεφάλαιο παρουσιάζονται δύο εφαρμογής. Η πρώτη εφαρμογή είναι μιας ουράς αναμονής σε ένα τηλεφωνικό κέντρο μιας τράπεζας στην οποία υπολογίζονται κάποιες βασικές παραμέτρους της θεωρίας των ουρών αναμονής και μεγαλύτερο ενδιαφέρον παρουσιάζει ο μέσος χρόνος που περιμένει ο πελάτης στην ουρά όταν υπάρχει ένας τηλεφωνητής και όταν υπάρχουν δύο. Η δεύτερη εφαρμογή ασχολείται με τα επείγοντα ενός νοσοκομείου όταν υπάρχουν τέσσερις και πέντε εφημερεύοντες ιατροί και γίνεται υπολογισμός των πόσων ασθενών χάνονται στην κάθε περίπτωση. Και στα δύο παραδείγματα γίνεται εφαρμογή της θεωρίας των κεφαλαίων δύο και τρία.

Στο κεφάλαιο πέντε γίνεται αναφορά στα συμπεράσματα που βγήκαν από το κεφάλαιο τέσσερα και τις εφαρμογές και δίνονται τα περιθώρια περεταίρω έρευνας πάνω στο θέμα των ουρών αναμονής και της μεταβλητότητας. Τέλος, στο τελευταίο κεφάλαιο αναφέρεται η βιβλιογραφία που χρησιμοποιήθηκε.

Πίνακας πινάκων

Πίνακας εικόνων

Κεφάλαιο 1: Θεωρία Ουρών Αναμονής

Ξεκινώντας με αυτό το κεφάλαιο αναλύεται η θεωρία των ουρών αναμονής σε συστήματα εξυπηρέτησης. Παρουσιάζονται μερικά από τα στοιχειώδη μοντέλα της θεωρίας ουρών αναμονής και δείχνεται το πώς αυτά βελτιώνουν την απόδοση σε ένα σύστημα εξυπηρέτησης. Τέλος, παρουσιάζονται κάποιες βασικές κατανομές και θεωρήματα όπως η εκθετική κατανομή που συναντάται συχνά σε προβλήματα ουρών.

1.1 Εισαγωγή

Η θεωρία ουρών (queuing theory) είναι ένας κλάδος των μαθηματικών και πιο συγκεκριμένα του τομέα της Επιχειρησιακής Έρευνας. Έχει την αρχή της σε μελέτες που έγιναν στις αρχές του $20^{ου}$ αιώνα από τον μαθηματικό A.K Erlang, ο οποίος πειραματίστηκε κυρίως πάνω σε ένα πρόβλημα συνωστισμού σε τηλεφωνικές γραμμές. Το πρόβλημα ήταν ότι σε ώρες αιχμής, ο εξυπηρετητής δεν προλάβαινε να απαντά σε όλα τα τηλεφωνήματα που έφθαναν με αποτέλεσμα να πρέπει οι πελάτες να περιμένουν. Αυτό που έκανε ο Erlang ήταν να προσπαθήσει να υπολογίσει την καθυστέρηση για έναν εξυπηρετητή. Η έρευνα συνεχίστηκε και αργότερα από τους Molina και Fry και μετά τον $2^{ο}$ Παγκόσμιο Πόλεμο η έρευνα αυτή επεκτάθηκε και σε άλλους τομείς.

Οι ουρές αναμονής είναι ένα φαινόμενο που συναντάμε όλοι στην καθημερινότητά μας και δημιουργούνται όταν η δυναμικότητα ενός συστήματος εξυπηρέτησης (service capacity) δεν επαρκεί για να ικανοποιήσει τη ζήτηση. Μπορούν όμως να υπάρξουν και σε περιπτώσεις όπου η δυναμικότητα είναι επαρκής. Ο λόγος που δημιουργούνται οι ουρές αναμονής είναι ότι οι πελάτες δεν φτάνουν με έναν σταθερό ρυθμό στο σύστημα, ούτε και εξυπηρετούνται όλοι στον ίδιο χρόνο. Οι πελάτες φτάνουν σε τυχαίες στιγμές και ο χρόνος εξυπηρέτησής τους ποικίλει, γεγονός που μας προκαλεί αβεβαιότητα για το αποτέλεσμα. Τα φαινόμενα που παρουσιάζουν αβεβαιότητα ονομάζονται στοχαστικά (stochastic) και αυτά είναι που μελετάμε κυρίως στη θεωρία ουρών αναμονής. Αυτά τα φαινόμενα λοιπόν είναι που οδηγούν στην αναμονή των πελατών και κατά συνέπεια η θεωρία ουρών έχει ως στόχο τη μείωση του χρόνου αναμονής των πελατών σε ένα σύστημα εξυπηρέτησης. Η μείωση του χρόνου αναμονής μπορεί να επιτευχθεί είτε με την αύξηση της δυναμικότητας του συστήματος, είτε με την βελτίωση του ρυθμού εξυπηρέτησης. Και στις δύο περιπτώσεις όμως, έχουμε κάποιο κόστος. Άρα πρέπει να βρεθεί μία χρυσή τομή μεταξύ της βελτίωσης του συστήματος εξυπηρέτησης και του κόστους που επιφέρει αυτή.

Όλα τα παραπάνω υπολογίζονται με κάποιους δείκτες απόδοσης. Πριν όμως προχωρήσουμε σε αυτούς θα ήταν καλό να δούμε τα βασικά στοιχεία ενός συστήματος εξυπηρέτησης. Αυτά είναι το πλήθος των πελατών, η διαδικασία αφίξεων, η πειθαρχία της ουράς και η εξυπηρέτηση, όπως φαίνεται στην Εικόνα 1.1

Εικόνα 1.1 Γενική μορφή ενός συστήματος εξυπηρέτησης (Πηγή: Tanner M. 1995. «Practical Queueing Analysis McGrawHil)

Ένα πλήθος πελατών εισέρχεται στο σύστημα. Οι πελάτες αυτοί φτάνουν με μία διαδικασία αφίξεων και αν δεν μπορέσουν να εξυπηρετηθούν άμεσα επειδή όλοι οι εξυπηρετητές είναι απασχολημένοι, θα περιμένουν στην ουρά μέχρι να έρθει η σειρά τους. Μόλις φτάσει η σειρά του πελάτη θα κάνει κάποιο χρόνο να εξυπηρετηθεί και με την περάτωση της εξυπηρέτησης ο πελάτης θα εξέλθει από το σύστημα.

1.2 Πλήθος πελατών

Το πλήθος των πελατών είναι το κύριο στοιχείο για ένα σύστημα εξυπηρέτησης. Το πλήθος αυτό μπορεί να χαρακτηριστεί άπειρο ή πεπερασμένο. Έστω για παράδειγμα ότι υπάρχει μία τράπεζα σε μία πόλη. Το πλήθος των πελατών σε αυτή την περίπτωση είναι το πλήθος των κατοίκων της πόλης. Παρόλο που ο αριθμός των κατοίκων της πόλης είναι πρακτικά πεπερασμένος, στην περίπτωσή που αναλύεται έστω ότι είναι άπειρος γιατί μιλάμε για ένα πολύ μεγάλο νούμερο και γιατί η εμφάνιση νέων πελατών δεν επηρεάζεται από των αριθμό εκείνων που εγκατέλειψαν το σύστημα.

Αντίστοιχα υπάρχει και το πεπερασμένο πλήθος όπου το πλήθος των πελατών που βρίσκεται στο σύστημα, επηρεάζει το ρυθμό αφίξεων νέων πελατών. Ένα τέτοιο παράδειγμα είναι οι μηχανές ενός εργοστασίου που χρειάζονται συντήρηση. Εδώ οι μηχανές είναι ο πελάτης και το συνεργείο είναι ο εξυπηρετητής. Έτσι λοιπόν ο αριθμός των πελατών είναι πεπερασμένος γιατί η άφιξη ενός νέου πελάτη (μηχανής) εξαρτάται από το πλήθος των πελατών (μηχανών) που βρίσκονται στο συνεργείο.

1.3 Διαδικασία αφίξεων

Ένα επίσης σημαντικό στοιχείο ενός συστήματος εξυπηρέτησης είναι η διαδικασία των αφίξεων των πελατών που έχει να κάνει με το ρυθμό που φτάνουν οι πελάτες στο σύστημα εξυπηρέτησης. Υπάρχουν δύο τύποι αφίξεων

- Κανονικές (ή προγραμματισμένες) αφίξεις
- Τυχαίες αφίξεις

Στις κανονικές αφίξεις, οι πελάτες φτάνουν ανά ίσα χρονικά διαστήματα, σαν να υπάρχει κάποιο πρόγραμμα. Αυτή η μορφή αφίξεων είναι η πιο απλή αλλά και αυτή που εμφανίζεται λιγότερο.

Αντίθετα οι τυχαίες αφίξεις συναντώνται πιο συχνά, αφού στις περισσότερες περιπτώσεις δεν γνωρίζεται η ακριβή ώρα άφιξης του πελάτη αλλά ούτε και αν θα εμφανιστεί μόνος του ή με άλλους μαζί.

Ένα από τα πρότυπα που χρησιμοποιούνται για να υπολογίσουμε τις τυχαίες αφίξεις, είναι αυτό της κατανομής Poisson, το οποίο χρησιμοποιείται πολύ σε διάφορες υπηρεσίες για να υπολογιστεί το διάστημα που μεσολαβεί μεταξύ δύο διαδοχικών αφίξεων. Μία ακόμα κατανομή που χρησιμοποιείται αρκετά είναι η εκθετική όπως θα δούμε και παρακάτω.

1.4 Η πειθαρχία της ουράς

Η πειθαρχία της ουράς πρακτικά έχει να κάνει με τους κανόνες προτεραιότητας που εξυπηρετούνται οι πελάτες. Αυτοί οι κανόνες είναι οι παρακάτω

Κανόνες προτεραιότητας εξαρτημένοι από το χρόνο επεξεργασίας

Κανόνες προτεραιότητας ανεξάρτητοι από το χρόνο επεξεργασίας

1. Κανόνες προτεραιότητας εξαρτημένοι από το χρόνο επεξεργασίας

   Ο πιο συχνός κανόνας προτεραιότητας που χρησιμοποιείται και σχετίζεται με το χρόνο εξυπηρέτησης είναι αυτός του συντομότερου χρόνου επεξεργασίας (shortest processing time SPT). Σύμφωνα με τον SPT, ο επόμενος διαθέσιμος εξυπηρετητής θα αναλάβει τον πελάτη με τον μικρότερο (αναμενόμενο) χρόνο εξυπηρέτησης σε σχέση με όλους τους άλλους πελάτες στην ουρά. Ο κανόνας SPT είναι πολύ αποτελεσματικός και αποδίδει καλά, με σεβασμό στον αναμενόμενο χρόνο αναμονής, καθώς επίσης και στην διακύμανση του χρόνου αναμονής.

2. Κανόνες προτεραιότητας ανεξάρτητοι από το χρόνο επεξεργασίας

   Ο πιο συχνά χρησιμοποιούμενος κανόνας προτεραιότητας βασίζεται στους χρόνους αφίξεων, αυτός δηλαδή που βρίσκεται πρώτος στην ουρά να εξυπηρετείται και πρώτος. Αυτή είναι η μέθοδος First In First Out (FIFO) ή αλλιώς First Comes First Served (FCFS) και είναι αυτή που παρατηρούμε σε όλες σχεδόν τις δημόσιες υπηρεσίες. Μία άλλη μέθοδος είναι η Last In First Out (LIFO) όπου εκεί ο τελευταίος που εισέρχεται στην ουρά εξυπηρετείται πρώτος. Για παράδειγμα στα μηνύματα στο κινητό πρώτο βλέπουμε αυτό που λάβαμε τελευταίο.

Εκτός από την χρήση των πληροφοριών της ώρας άφιξης, πολλές καταστάσεις στην πράξη απαιτούν χαρακτηριστικά, όπως η σημασία ή το επείγον της υπόθεσης, που εξετάζεται στον κανόνα προτεραιότητας, όπως για παράδειγμα στα επείγοντα ενός νοσοκομείου. Εδώ υπάρχουν δύο περιπτώσεις προτεραιότητας. Η απλή προτεραιότητα ή προτεραιότητα χωρίς διακοπή (non preemptive), όπου επιλέγεται ο πελάτης με την υψηλότερη προτεραιότητα και την απόλυτη προτεραιότητα ή προτεραιότητα με διακοπή (preemptive), στην οποία αν εξυπηρετείται κάποιος πελάτης χαμηλής προτεραιότητας και εισέλθει στην ουρά κάποιος υψηλής προτεραιότητας, τότε σταματάει η εξυπηρέτηση του πρώτου ώστε να εξυπηρετηθεί ο δεύτερος.

Κάτι άλλο εξίσου σημαντικό για την ουρά αναμονής είναι το μέγεθός της το οποίο είναι είτε πεπερασμένο είτε άπειρο. Μια άπειρη ουρά μπορεί να είναι οποιουδήποτε μεγέθους χωρίς άνω όριο όπως για παράδειγμα η ουρά σε έναν κινηματογράφο. Από την άλλη μεριά, η πεπερασμένη ουρά έχει συγκεκριμένο μέγεθος όπως για παράδειγμα σε ένα συνεργείο αυτοκινήτων που θα πρέπει να επισκευαστούν τα αυτοκίνητα που υπάρχουν για να έρθουν άλλα καινούρια. Η συμπεριφορά λοιπόν των πελατών εξαρτάται άμεσα από τη χωρητικότητα της ουράς. Για παράδειγμα, αν η ουρά είναι πολύ μεγάλη μπορεί οι πελάτης να μην μπουν στην διαδικασία να περιμένουν (balking) ή ακόμα και να περιμένουν στην ουρά, μπορεί η αναμονή να είναι τόσο μεγάλη που να βαρεθούν και να φύγουν (reneging). Αυτό επηρεάζει άμεσα το ρυθμό αφίξεων των πελατών στο σύστημα ο οποίος μηδενίζεται αφού λόγω μικρής χωρητικότητας της ουράς οι πελάτες είτε δεν εισέρχονται είτε την εγκαταλείπουν.

1.5 Εξυπηρέτηση

Τέλος, για τη διαδικασία εξυπηρέτησης έχουμε κάποια βασικά χαρακτηριστικά που μας ενδιαφέρουν. Αυτά είναι ο αριθμός των θέσεων εξυπηρέτησης, οι φάσεις της εξυπηρέτησης και κυρίως ο χρόνος που χρειάζεται για να εξυπηρετηθεί ένας πελάτης.

Ο αριθμός των θέσεων εξυπηρέτησης είναι ο αριθμός των παράλληλων σημείων που μπορούν να εξυπηρετούν τους πελάτες. Μπορεί να υπάρχει μία θέση εξυπηρέτησης, μπορεί όμως και πολλές θέσεις που να δουλεύουν παράλληλα ακόμα και για να εξυπηρετήσουν μία μόνο ουρά. Για παράδειγμα στα ταμεία μιας τράπεζας ενώ η ουρά των πελατών είναι μία, οι υπάλληλοι που εξυπηρετούν μπορεί να είναι περισσότεροι.

Σε κάποιες περιπτώσεις ο πελάτης μπορεί να εξυπηρετηθεί από μία θέση εξυπηρέτησης και να έχει ολοκληρώσει τη δουλειά του. Σε αυτή την περίπτωση η φάση της εξυπηρέτησης είναι μία. Υπάρχουν όμως και περιπτώσεις όπου ο πελάτης πρέπει να περάσει από περισσότερες από μία φάσεις για να ολοκληρωθεί η εξυπηρέτησή του. Όπως για παράδειγμα στο αεροδρόμιο πρέπει πρώτα να γίνει έλεγχος των αποσκευών, μετά έλεγχος του εισιτηρίου και των προσωπικών εγγράφων και τέλος να γίνει η επιβίβαση στο αεροπλάνο.

Τελευταίο και πιο σημαντικό στοιχείο στην εξυπηρέτηση των πελατών είναι ο χρόνος εξυπηρέτησης, τα άτομα δηλαδή που μπορεί να εξυπηρετήσει ένας εξυπηρετητής στη μονάδα του χρόνου. Όπως και στην περίπτωση των αφίξεων έτσι και εδώ μπορούμε να υποθέσουμε ότι αυτός ο ρυθμός εξυπηρέτησης είναι καθορισμένος. Στις περισσότερες όμως περιπτώσεις θα πρέπει να χρησιμοποιήσουμε κάποιο μαθηματικό μοντέλο για να μπορέσουμε να τον υπολογίσουμε. Η πιο συχνή κατανομή που χρησιμοποιείται για τον υπολογισμό του χρόνου εξυπηρέτησης είναι η εκθετική κατανομή.

Τελειώνοντας τη διαδικασία της εξυπηρέτησης, ο πελάτης είναι πιθανόν να επιστρέψει στην πηγή. Υπάρχουν όμως και περιπτώσεις όπου ο πελάτης θα επιστρέψει καθυστερημένα στην πηγή ή δεν θα επιστρέψει καθόλου. Καλό θα είναι ο χρήστης να μην ξεχνά να αξιολογεί και αυτό πριν αποφασίσει να χρησιμοποιήσει κάποιο μοντέλο.

1.6 Συμβολισμός για τα συστήματα ουρών (Kendal Notation)

Ο D.G Kendall (1953) παρέθεσε ένα συμβολισμό για τα συστήματα ουρών ώστε να βοηθήσει μελλοντικούς ερευνητές και να υπάρχει μία κοινή γλώσσα αναφοράς. Ο συμβολισμός αυτός ξεκίνησε με τα τρία γράμματα A/S/c όπου A ο χρόνος μεταξύ των διαδοχικών αφίξεων, S η κατανομή του χρόνου εξυπηρέτησης και c ο αριθμός των εξυπηρετητών. Με τα χρόνια ο συμβολισμός επεκτάθηκε και καταλήξαμε σήμερα να χρησιμοποιούμε τα εξής έξι γράμματα «A/S/c/K/N/D» όπου K η χωρητικότητα της ουράς, N το πλήθος των πελατών στην πηγή και D η πειθαρχία της ουράς.

Πίνακας 1.1Ανάλυση συμβόλων (Πηγή: Π-.Χ.Γ. Βασιλείου «Στοχαστικές μέθοδοι στις επιχειρησιακές έρευνες»)

Χαρακτηριστικά	**Σύμβολο**	**Ερμηνεία**
Χρόνος μεταξύ αφίξεων (A)	M D E_k G	Κατανομή Poisson Κανονική κατανομή Κατανομή Erlang Γενική κατανομή
Κατανομή χρόνου εξυπηρέτησης (S)	M D E_k G	Κατανομή Poisson Κανονική κατανομή Κατανομή Erlang Γενική κατανομή
Αριθμός εξυπηρετητών (c)	1, 2, 3,…..,∞	
Χωρητικότητα της ουράς (K)	1, 2, 3,…..,∞	
Πληθυσμός πελατών (N)	1, 2, 3,…..,∞	
Πειθαρχία ουράς (D)	FIFO LIFO SIRO PNPN	Οι πελάτες εξυπηρετούνται με τη σειρά που έχουν φτάσει Ο τελευταίος που έφτασε εξυπηρετείται πρώτος Τυχαία εξυπηρέτηση Εξυπηρέτηση με προτεραιότητα

1.7 Μοντέλα ουρών αναμονής

1.7.1 Το μοντέλο M/M/1

Έστω ότι στο βασικό σύστημα ουράς M/M/1 οι αφίξεις είναι σταθερές και ακολουθούν την κατανομή Poisson με παράμετρο λ. Υπάρχει ένα σημείο εξυπηρέτησης και ο ρυθμός εξυπηρέτησης είναι σταθερός και ακολουθεί την εκθετική κατανομή με μέση τιμή μ. Τέλος, η ουρά που θα δημιουργηθεί θα είναι αόριστη γιατί θα έχουμε λ <μ.

Θα υπολογιστεί τώρα το $p_n(t)$, που είναι η πιθανότητα να βρίσκεται η ουρά σε μία συγκεκριμένη κατάσταση n στη μονάδα του χρόνου t.

$$\sum_{n=0}^{\infty}(\frac{\lambda}{\mu})^n = \frac{1}{1-(\frac{\lambda}{\mu})}, \tag{1.1}$$

Και

$$\sum_{n=0}^{\infty} P_n = 1 \tag{1.2}$$

$$\sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty}(\frac{\lambda}{\mu})^n P_0 = \frac{P_0}{1-\frac{\lambda}{\mu}} = 1 \tag{1.3}$$

Άρα

$$P_0 = 1 - \frac{\lambda}{\mu} \tag{1.4}$$

$$P_n = \left(\frac{\lambda}{\mu}\right)^n (1 - \frac{\lambda}{\mu}) \quad \text{n=1,2,…..} \tag{1.5}$$

Γνωρίζοντας ότι λ είναι ο μέσος ρυθμός άφιξης και μ ο μέσος ρυθμός εξυπηρέτησης, είναι δυνατόν να υπολογιστούν και άλλα μεγέθη τα οποία μας βοηθούν στην καλύτερη κατανόηση αλλά και επίλυση των προβλημάτων αναμονής. Τέτοια μεγέθη είναι :

$L = \frac{\lambda}{\mu-\lambda} = ο\ μέσος\ αριθμός\ πελατών\ στο\ σύστημα$

$L_q = \frac{\lambda^2}{\mu(\mu-\lambda)} = μέσος\ αριθμός\ πελατών\ στην\ ουρά$

$L_w = \frac{\mu}{\mu-\lambda}$
$= μέσος\ αριθμός\ πελατών\ σε\ ουρές\ που\ υπάρχει\ πελάτης\ που\ προηγείται$

$W = \frac{1}{\mu-\lambda} = μέσος\ χρόνος\ που\ περνά\ ένας\ πελάτης\ στο\ σύστημα$

$$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = μέσος\ χρόνος\ που\ περνά\ ένας\ πελάτης\ στην\ ουρά\ αναμονής$$

$$P(n > k) = (\frac{\lambda}{\mu})^{k+1}$$
$$= πιθανότητα\ ότι\ περισσότεροι\ από\ k\ πελάτες\ είναι\ στο\ σύστημα$$

1.7.2 Το μοντέλο M/M/1/K. (Πεπερασμένης χωρητικότητας)

Εδώ φαίνεται μία ουρά με τα χαρακτηριστικά της προηγούμενης, με τη μόνη διαφορά ότι εδώ υπάρχει συγκεκριμένη χωρητικότητα K. Θα ισχύουν οι εξισώσεις της προηγούμενης παραγράφου αλλά για n<K. Για n=K

$$p_k = \frac{\lambda}{\mu} p_{k-1} \tag{1.6}$$

Το σύστημα λοιπόν θα πάρει την παρακάτω μορφή

$$p_1 = \frac{\lambda}{\mu} p_0$$

$$p_{n+1} = \frac{\lambda + \mu}{\mu} p_n - \frac{\lambda}{\mu} p_{n-1} \tag{1.7}$$

Από την M/M/1

$$p_n = \rho^n p_0\ όπου\ \rho = \frac{\lambda}{\mu}$$

Εδώ όμως υπάρχει και μία προϋπόθεση

$$\sum_{n=0}^{k} p_n = 1$$

Άρα

$$p_0 = \frac{1}{\sum_{n=0}^{k} p_n}$$

Όπου

$$\sum_{n=0}^{k} \rho^n = \begin{cases} \frac{1 - \rho^{k+1}}{1 - \rho}, & \rho \neq 1 \\ k + 1, & \rho = 1 \end{cases} \tag{1.8}$$

Και βγαίνει το εξής

$$p_n = \begin{cases} \frac{(1-\rho)\rho^n}{1-\rho^{k+1}}, \rho \neq 1 \\ \frac{1}{1+k}, \rho = 1 \end{cases} \tag{1.9}$$

1.7.3 Το μοντέλο M/M/S

Έστω σε αυτό το μοντέλο ότι ο αριθμός των εξυπηρετητών είναι συγκεκριμένος έστω S και κάθε πελάτης που φτάνει πηγαίνει σε αυτόν που δεν είναι απασχολημένος. Όπως και στις προηγούμενες παραγράφους η διαδικασία άφιξης είναι Poisson με παράμετρο λ και χρόνο εξυπηρέτησης που ακολουθεί την εκθετική κατανομή με παράμετρο μ.

Αν οι πελάτες είναι λιγότεροι από S τότε η μέση τιμή αναχωρήσεων είναι nμ αφού είναι απασχολημένοι n<S, ενώ αν υπάρχουν περισσότεροι από S πελάτες η μέση τιμή αναχωρήσεων είναι Sμ. Και έτσι

$$\mu_n = \begin{cases} n\mu, & 1 \leq n < S \\ S\mu, & n \geq S \end{cases}$$

Επειδή η διαδικασία αυτή είναι μία διαδικασία «γέννησης- θανάτου» όπου εκεί

$$p_n = p_0 \prod_{i=1}^{n} \frac{\lambda_{i-1}}{\mu_i}$$

Χρησιμοποιώντας αυτή τη σχέση βγαίνει ο τύπος»

$$p_n = \begin{cases} \frac{\lambda^n}{n!\,\mu^n} p_0, & 1 \leq n < S \\ \frac{\lambda^n}{S^{n-S} S!\,\mu^n} p_0, & n \geq S \end{cases} \tag{1.10}$$

Έχοντας όπως και πριν τη συνθήκη $\sum_{n=0}^{\infty} p_n = 1$ βρίσκεται το p_0

$$p_0 = \left\{ \sum_{n=0}^{S-1} \frac{1}{n!} (\frac{\lambda}{\mu})^n + \frac{1}{S!} (\frac{\lambda}{\mu})^S (\frac{S\mu}{S\mu - \lambda}) \right\}^{-1} \tag{1.11}$$

Έχοντας υπολογίσει το p_0 μπορεί τώρα να υπολογίσουμε και το μήκος της ουράς το οποίο εξ' ορισμού είναι

$$L_q = \sum_{n=S}^{\infty} \frac{n}{S^{n-S} S!} (\frac{\lambda}{\mu})^n p_0 - \sum_{n=0}^{\infty} \frac{S}{S^{n-S} S!} (\frac{\lambda}{\mu})^n p_0$$

Και αναλύοντάς το καταλήγουμε στο εξής

$$L_q = \frac{(\frac{\lambda}{\mu})^S \lambda\mu}{(S-1)!\,(S\mu-\lambda)^2} p_0 \tag{1.12}$$

Για τον υπολογισμό του L όπου είναι η μέση τιμή του αριθμού των πελατών και όντας γνωστό ότι L=λW:

$$L = \frac{\lambda}{\mu} + \frac{(\frac{\lambda}{\mu})^S \lambda\mu}{(S-1)!\,(S\mu-\lambda)} p_0 \tag{1.13}$$

Τέλος, υπολογίζονται τα W_q και W όπου είναι γνωστό από πριν ότι $W_q = \frac{L_q}{\lambda}$ και $W = W_q + \frac{1}{\mu}$ και έτσι

$$W_q = \frac{L_q}{\lambda} = \frac{(\frac{\lambda}{\mu})^S \mu}{(S-1)!\,(S\mu-\lambda)^2} p_0 \tag{1.14}$$

Και

$$W = \frac{1}{\mu} + \frac{(\frac{\lambda}{\mu})^S \mu}{(S-1)!\,(S\mu-\lambda)^2} p_0 \tag{1.15}$$

1.7.4 Το μοντέλο M/M/S/K (Πεπερασμένη χωρητικότητα)

Σε αυτήν την περίπτωση αναλύεται την ουρά M/M/S όμως εδώ δεν μπορούν να υπάρχουν παραπάνω από K πελάτες. Αυτό αλλάζει τα λ και μ τα οποία γίνονται

$$\lambda_n = \begin{cases} \lambda, & 0 \le n \le K \\ 0, & n \ge K \end{cases}$$

$$\mu_n = \begin{cases} n\mu, & 0 \le n \le S \\ S\mu, & S \le n \le K \end{cases}$$

Έχοντας από πριν τις εξισώσεις

$$p_n = p_0 \prod_{i=1}^{n} \frac{\lambda_{i-1}}{\mu_i}$$

Και

$$p_n = \begin{cases} \dfrac{\lambda^n}{n!\,\mu^n} p_0, & 1 \le n < S \\ \dfrac{\lambda^n}{S^{n-S} S!\,\mu^n} p_0, & n \ge S \end{cases}$$

Είναι

$$p_n = \begin{cases} \dfrac{\lambda^n}{n!\,\mu^n} p_0, & 1 \le n < S \\ \dfrac{\lambda^n}{S^{n-S} S!\,\mu^n} p_0, & S \le n \le K \end{cases} \tag{1.16}$$

Γνωρίζοντας ότι $\sum_{n=0}^{\infty} p_n = 1$ συνεπάγεται

$$p_0 = \left\{ \sum_{n=0}^{S-1} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n + \frac{1}{S!} \left(\frac{\lambda}{\mu}\right)^S \left(\frac{S\mu}{S\mu - \lambda}\right) \right\}^{-1} \tag{1.17}$$

Αλλά και

$$\sum_{n=0}^{\infty} \frac{1}{S^{n-S} S!} \left(\frac{\lambda}{\mu}\right)^n = \begin{cases} \dfrac{\left(\frac{\lambda}{\mu}\right)^S}{S!} \dfrac{\left[1 - \left(\frac{\lambda}{\mu S}\right)^{K-S+1}\right]}{1 - \frac{\lambda}{S\mu}}, & \dfrac{\lambda}{S\mu} \ne 1 \\ \dfrac{\left(\frac{\lambda}{\mu}\right)^2}{S!} (K - S + 1), & \dfrac{\lambda}{S\mu} = 1 \end{cases} \tag{1.18}$$

Και τελικά

$$p_0 = \begin{cases} \left\{ \sum_{n=0}^{S-1} \frac{1}{n!} (\frac{\lambda}{\mu})^n + \frac{(\frac{\lambda}{\mu})^S}{S!} \frac{\left[1 - \left(\frac{\lambda}{\mu S} \right)^{K-S+1} \right]}{1 - \frac{\lambda}{S\mu}} \right\}^{-1} & , \frac{\lambda}{S\mu} \neq 1 \\ \left\{ \sum_{n=0}^{S-1} \frac{1}{n!} (\frac{\lambda}{\mu})^n + \frac{\left(\frac{\lambda}{\mu}\right)^2}{S!} (K - S + 1) \right\}^{-1} & , \frac{\lambda}{S\mu} = 1 \end{cases} \tag{1.19}$$

Και τέλος υπολογίζεται το μήκος της ουράς L_q, το μέγεθος του συστήματος L και τα W και W_q.

$$L_q = \frac{p_0 (S \frac{\lambda}{S\mu})^S \frac{\lambda}{S\mu}}{S!} \frac{\lambda}{S\mu} \left\{ 1 - \left(\frac{\lambda}{S\mu} \right)^{K-S+1} - (1 - \frac{\lambda}{S\mu})(K - S + 1)(\frac{\lambda}{S\mu})^{K-S} \right\} \tag{1.20}$$

$$L = \frac{p_0 (S \frac{\lambda}{S\mu})^S \frac{\lambda}{S\mu}}{S!} \frac{\lambda}{S\mu} \left\{ 1 - \left(\frac{\lambda}{S\mu} \right)^{K-S+1} - (1 - \frac{\lambda}{S\mu})(K - S + 1)(\frac{\lambda}{S\mu})^{K-S} \right\} + S - p_0 \sum_{n=0}^{S-1} \frac{(S-n)(\frac{\lambda}{\mu})^n}{n!} \tag{1.21}$$

Αφού L=Wλ΄ όπου λ΄=λ(1-p_k).Αντίστοιχα υπολογίζεται και τα $W_q = \frac{L_q}{\lambda'}$ και $W = W_q + \frac{1}{\mu}$

1.8 Κατανομές και θεωρίες

1.8.1 Κατανομή Poisson

Η κατανομή Poisson είναι η κατανομή των σπάνιων γεγονότων και χρησιμοποιείται όταν πρέπει να μετρηθεί ο αριθμός των εμφανίσεων στη μονάδα του χρόνου. Η τυχαία μεταβλητή X έχει συνάρτηση πυκνότητας πιθανότητας τον τύπο:

$$P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!} \tag{1.22}$$

Η Poisson χρησιμοποιείται σε αρκετά προβλήματα όπως για να βρεθεί ο αριθμός των τηλεφωνικών κλήσεων ενός τηλεφωνικού κέντρου σε ένα συγκεκριμένο χρονικό διάστημα, τον αριθμό των τυπογραφικών λαθών σε ένα κείμενο και άλλα. Γενικά εμφανίζεται όταν έχουμε έναν αριθμό παρατηρήσεων σε ένα χρονικό διάστημα.

Αυτή η κατανομή χρησιμοποιείται πολύ και στη θεωρία ουρών και κυρίως στη διαδικασία αφίξεων με τύπο:

$$p_{ij}(s,t) = prob\{X(t) = j|X(s) = i\}$$

Όπου X(t) τυχαία μεταβλητή που δηλώνει τον αριθμό των αφίξεων πελατών σε ένα χρονικό διάστημα (0,t) και έχουμε s<t.

Ας δούμε τώρα τη διαδικασία αφίξεων να γίνεται με τον εξής τρόπο

1. Αφίξεις που εμφανίζονται σε ξένα μεταξύ τους διαστήματα είναι ανεξάρτητες.
2. Για πολύ μικρό Δt, υπάρχει σταθερή λ τέτοια ώστε η πιθανότητα να γίνει άφιξη στο διάστημα (t,t+Δt] να δίνεται από τις σχέσεις:
   a) $p_{ii}(t,\Delta t) = 1 - \lambda\Delta t + 0(\Delta t)$
   b) $p_{i,i+1}(t,\Delta t + t) = \lambda\Delta t + 0(\Delta t)$
   c) $\sum_{j=i+2}^{\infty} p_{ij}(t,t+\Delta t) = 0(\Delta t)$
   d) $p_{ij}(t,t+\Delta t) = 0$

Όπου 0(Δt) περιλαμβάνει όλους τους όρους που τείνουν στο 0 πολύ πιο γρήγορα από το Δt.

Αποδεικνύεται πως ισχύει το παρακάτω θεώρημα

Θεώρημα 1.1

Έστω X(t) νια τυχαία μεταβλητή σύμφωνα με παραπάνω, τότε

$$p_{on}(0,t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \tag{1.23}$$

Αυτός ο τύπος δείχνει πως για ένα χρονικό διάστημα (0,t) η X ακολουθεί κατανομή Poisson με παράμετρο λt

1.8.2 Εκθετική κατανομή

Πρόκειται για συνεχή κατανομή με τυχαία μεταβλητή X και συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \begin{cases} \mu e^{-\mu x}, x \geq 0 \\ 0, \alpha\lambda\lambda o\acute{\upsilon} \end{cases}$$

Και ισχύει $E(x) = \frac{1}{\mu}$ και $\sigma^2(x) = \frac{1}{\mu^2}$

Η εκθετική κατανομή συνδέεται άμεσα με την κατανομή Poisson άρα και με τη θεωρία ουρών. Αυτό θα δείξουμε με το παρακάτω θεώρημα

Θεώρημα 1.2

Οι τυχαίες μεταβλητές $T_1, T_2, \ldots, T_n$ που εκφράζουν το χρόνο ανάμεσα σε δύο διαδοχικές αφίξεις σε μία Poisson διαδικασία αφίξεων είναι ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την εκθετική κατανομή, δηλ,

$prob(T_n \leq x) = 1 - e^{-\lambda x}$

Τέλος, μία σημαντική ιδιότητα μιας τυχαίας εκθετικής μεταβλητής X με παράμετρο μ είναι η ιδιότητα της αμνησίας. Αυτή η ιδιότητα μας λέει ότι για όλα τα x≥0 και τα t≥0

$$P(X > x + t | X > t) = P(X > x) = e^{-\mu x} \qquad (1.24)$$

Έτσι η εναπομένουσα διάρκεια ζωής της X κατανέμεται και πάλι εκθετικά με μέσο όρο 1/μ.

1.8.3 Νόμος του Little

Σε ένα σύστημα αναμονής υπάρχουν κάποια αντικείμενα τα οποία μπορεί να είναι στην ουρά ή να εξυπηρετούνται ή ακόμα και κάποια να είναι στην ουρά και κάποια να εξυπηρετούνται. Η ερμηνεία θα εξαρτηθεί από την εφαρμογή ή τους στόχους του ατόμου που δημιουργεί το μοντέλο. Για παράδειγμα έστω ένα κελάρι με κρασιά. Έστω ότι ένα μπουκάλι κρασί εισάγεται στο σύστημα όταν μπαίνει στο κελάρι. Κάθε μπουκάλι μένει στο σύστημα μέχρι να αποφασίσει ο ιδιοκτήτης να το καταναλώσει. Αν παρατηρηθεί το ράφι του κρασιού σαν ένας εξυπηρετητή ενός καναλιού, ο χρόνος εξυπηρέτησης είναι ο χρόνος μεταξύ των διαδοχικών "μετακομίσεων". Είναι ενδιαφέρον να σημειωθεί ότι δεν είναι γνωστό ποιο μπουκάλι θα πάρει πρώτα ο ιδιοκτήτης και ότι δεν θα ακολουθήσει απαραίτητα τον κανόνα FIFO. Σε κάθε περίπτωση, για να υπολογιστεί ο μέσος αριθμός φιαλών στο κελάρι ή ο μέσος χρόνος που περνάει κάθε μπουκάλι στο κελάρι πρέπει να παρατηρηθεί όλο το σύστημα να αποτελείται από ουρά και από εξυπηρέτηση.

Ο νόμος του Little λέει ότι, υπό συνθήκες σταθερής κατάστασης, ο μέσος αριθμός των στοιχείων σε ένα σύστημα ουρών αναμονής ισούται με το μέσο ρυθμό με τον οποίο φτάνουν τα αντικείμενα πολλαπλασιαζόμενο με το μέσο χρόνο που μένει ένα αντικείμενο στο σύστημα. Έχουμε:

L= μέσος αριθμός αντικειμένων στο σύστημα αναμονής

W=μέσος όρος αναμονής στο σύστημα για ένα αντικείμενο

λ= μέσος αριθμός αντικειμένων που φτάνουν ανά μονάδα χρόνου

$$L = W * \lambda \qquad (1.25)$$

Η σχέση αυτή είναι πολύ χρήσιμη λόγο της γενικότητάς της. Χρησιμοποιείται πολύ στην ουρά G/G/K. Επιπλέον ισχύουν οι σχέσεις:

$$L_q = \lambda * W_q$$

$$L = L_q + (\frac{\lambda}{\mu})$$

$$W_q = \frac{L}{\mu}$$

Όπου Q ο αριθμός των πελατών στο σύστημα σε κατάσταση στατικής ισορροπίας.

1.8.4 Ο πίνακας συνάρτησης απώλειας του Erlang

Ο πίνακας της συνάρτηση απώλειας του Erlang περιέχει την πιθανότητα όπου στο στάδιο μιας διαδικασίας με m παράλληλες πηγές που περιέχουν m μονάδες ροής, όλες οι m πηγές χρησιμοποιούνται. Τα διαστήματα μεταξύ των αφίξεων των μονάδων ροής είναι εκθετικά κατανεμημένα με μέση τιμή α και οι χρόνοι εξυπηρέτησης, που δεν ακολουθούν απαραίτητα εκθετική κατανομή, έχουν μέση τιμή p.

Μη έχοντας χώρο αναμονής, όταν φτάνει μία μονάδα και όλοι οι εξυπηρετητές είναι απασχολημένοι, τότε η μονάδα εγκαταλείπει το σύστημα χωρίς να έχει εξυπηρετηθεί. Οι στήλες του πίνακα αντιπροσωπεύουν τον αριθμό των πηγών, m, στο σύστημα και οι στήλες είναι η αναλογία του μέσου χρόνου εξυπηρέτησης προς το μέσο χρόνο μεταξύ των αφίξεων, r=p/α.

Σύμφωνα λοιπόν με τα παραπάνω, η πιθανότητα να είναι όλοι οι εξυπηρετητές απασχολημένοι, $P_m(r)$ υπολογίζεται με τον παρακάτω τύπο:

$$Probability \,\{όλοι\ οι\ εξυπηρετητές\ m\ να\ είναι\ απασχολημένοι\} \\ = P_m(r) = \frac{\frac{r^m}{m!}}{1 + \frac{r}{1!} + \frac{r^2}{2!} + \cdots . + \frac{r^m}{m!}} \qquad (1.26)$$

Υπάρχει όμως και πιο απλός τρόπος υπολογισμού αυτής της πιθανότητα και είναι με τη χρήση του πίνακα που αναφέρθηκε παραπάνω.

Πίνακας 1.2 Πίνακας απώλειας του Erlang(Πηγή: Cachon- Terwiesch “Matching Supply with Demand”)

		m									
		1	2	3	4	5	6	7	8	9	10
	0,1	0,0909	0,0045	0,0002	0	0	0	0	0	0	0
	0,2	0,1667	0,0164	0,0011	0,0001	0	0	0	0	0	0
	0,25	0,2	0,0244	0,002	0,0001	0	0	0	0	0	0
	0,3	0,2308	0,0335	0,0033	0,0003	0	0	0	0	0	0
	0,33	0,25	0,04	0,0044	0,0004	0	0	0	0	0	0
	0,4	0,2857	0,0541	0,0072	0,0007	0,0001	0	0	0	0	0
	0,5	0,3333	0,0769	0,0127	0,0016	0,0002	0	0	0	0	0
	0,6	0,375	0,1011	0,0198	0,003	0,0004	0	0	0	0	0
	0,67	0,4	0,1176	0,0255	0,0042	0,0006	0,0001	0	0	0	0
	0,7	0,4118	0,126	0,0286	0,005	0,0007	0,0001	0	0	0	0
	0,75	0,4286	0,1385	0,0335	0,0062	0,0009	0,0001	0	0	0	0
	0,8	0,4444	0,1509	0,0387	0,0077	0,0012	0,0002	0	0	0	0
r	0,9	0,4737	0,1757	0,0501	0,0111	0,002	0,0003	0	0	0	0
	1	0,5	0,2	0,0625	0,0154	0,0031	0,0005	0,0001	0	0	0
	1,1	0,5238	0,2237	0,0758	0,0204	0,0045	0,0008	0,0001	0	0	0
	1,2	0,5455	0,2466	0,0898	0,0262	0,0063	0,0012	0,0002	0	0	0
	1,25	0,5556	0,2577	0,097	0,0294	0,0073	0,0015	0,0003	0	0	0
	1,3	0,5652	0,2687	0,1043	0,0328	0,0085	0,0018	0,0003	0,0001	0	0
	1,33	0,5714	0,2759	0,1092	0,0351	0,0093	0,0021	0,0004	0,0001	0	0
	1,4	0,5833	0,2899	0,1192	4	0,0111	0,0026	0,0005	0,0001	0	0
	1,5	0,6	0,3103	0,1343	0,048	0,0142	0,0035	0,0008	0,0001	0	0
	1,6	0,6154	0,3299	0,1496	0,0565	0,0177	0,0047	0,0011	0,0002	0	0
	1,67	0,625	0,3425	0,1598	0,0624	0,0204	0,0056	0,0013	0,0003	0,0001	0
	1,7	0,6296	0,3486	0,165	0,0655	0,0218	0,0061	0,0015	0,0003	0,0001	0
	1,75	0,6364	0,3577	0,1726	0,0702	0,024	0,0069	0,0017	0,0004	0,0001	0
	1,8	0,6429	0,3665	0,1803	0,075	0,0263	0,0078	0,002	0,0005	0,0001	0
	1,9	0,6552	0,3836	0,1955	0,085	0,0313	0,0098	0,0027	0,0006	0,0001	0
	2	0,6667	0,4	0,2105	0,0952	0,0367	0,0121	0,0034	0,0009	0,0002	0

Έστω ένα σύστημα με 4 παράλληλες πηγές. Μία μονάδα ροής να φτάνει κάθε α=3 λεπτά και χρειάζεται p=2 λεπτά να εξυπηρετηθεί. Έχουμε ότι r=2/3=0.67. Άρα πηγαίνοντας στην 4η στήλη και την γραμμή όπου βρίσκεται το 0,67 υπολογίζεται ότι $P_m(r)$=0.0042.

Κεφάλαιο 2: Προβλήματα ουρών αναμονής με μεταβλητότητα

Το παρόν κεφάλαιο προέρχεται κυρίως από το βιβλίο Matching Supply with Demand των Cachon- Terwiesch και το κεφάλαιο 8 και ασχολείται με τη μεταβλητότητα και κυρίως στη διαδικασία των αφίξεων. Υπολογίζεται ο μέσος χρόνος αφίξεων με έναν και με πολλούς εξυπηρετητές και αναφέρει τα επίπεδα εξυπηρέτησης στα προβλήματα των χρόνων αναμονής. Τέλος, παρουσιάζονται οι οικονομικές επιπτώσεις και οι επιπτώσεις συγκέντρωσης και προτείνονται τρόποι μείωσης της μεταβλητότητας.

2.1 Εισαγωγή

Όλοι οι καταναλωτές περνάν αρκετό χρόνο περιμένοντας σε διάφορες ουρές. Αυτές ουρές μπορεί να είναι είτε φυσικές (σούπερ μάρκετ) είτε εικονικές (η αναμονή σε ένα τηλεφωνικό κέντρο). Ο χρόνος αναμονής λοιπόν είναι ένα βασικό στοιχείο στη ζωή μας αλλά και εξίσου ενοχλητικό. Θα πρέπει να γίνει γνωστό ότι υπάρχουν δύο τύποι χρόνου αναμονής οι οποίοι δεν είναι τόσο σημαντικοί για τους καταναλωτές, έχουν όμως μεγάλη σημασία στη διοίκηση λειτουργιών.

Αυτοί οι δύο τύποι είναι οι εξής:

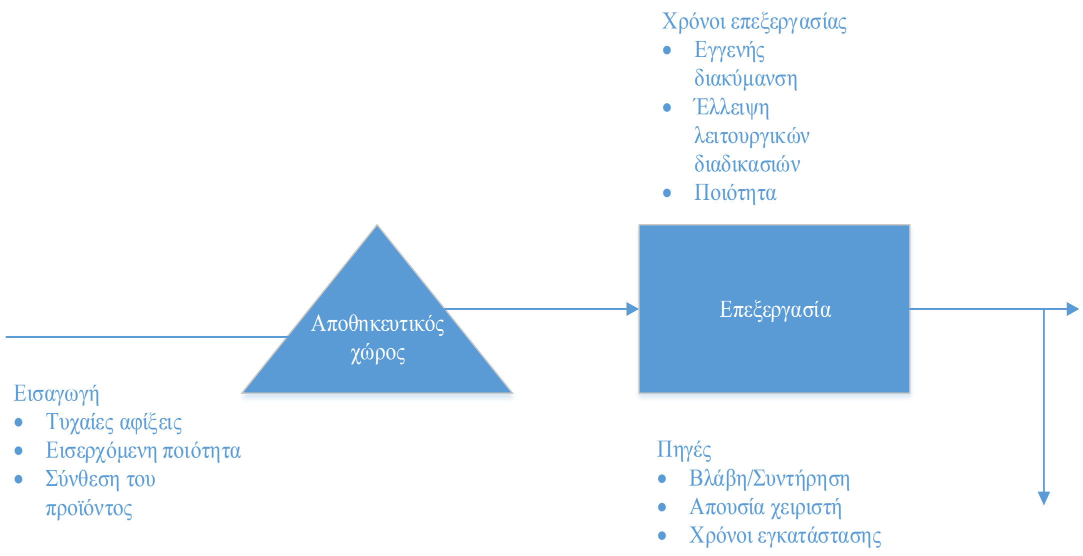
- Ο πρώτος τύπος είναι όταν ο αναμενόμενος δείκτης ζήτησης ξεπερνά τον αναμενόμενο δείκτη προμήθειας για ένα περιορισμένο χρονικό διάστημα. Αυτό συμβαίνει κυρίως σε περιπτώσεις σταθερών επιπέδων χωρητικότητας και ζήτησης που εμφανίζονται εποχιακά. Αυτό οδηγεί πάνω από 100% αξιοποίηση σε κάποια χρονική περίοδο. Οι σειρές που δημιουργούνται στην πύλη ενός αεροδρομίου μετά την ανακοίνωση μιας πτήσης είναι ένα παράδειγμα τέτοιων ουρών.
- Ο δεύτερος τύπος έχει να κάνει με την ύπαρξη της μεταβλητότητας. Και σε αυτήν την περίπτωση αν υπάρχει αξιοποίηση μικρότερη του 100%, προκύπτουν ουρές. Τέτοιες ουρές μπορεί να αποδοθούν πλήρως στην ύπαρξη της μεταβλητότητας, καθώς υπάρχει, κατά μέσο όρο, αρκετή χωρητικότητα για την υπάρχουσα ζήτηση.

  Όπως μπορεί να γίνει αντιληπτό στον πρώτο τύπο η αιτία του προβλήματος είναι η χωρητικότητα, ενώ στον δεύτερο είναι η μεταβλητότητα. Η μεταβλητότητα κάνει ο χρόνο αναμονής απρόβλεπτο για όλους. Κάποιες φορές περιμένει ο πελάτης (ζήτηση) να εξυπηρετηθεί (προμήθεια) και κάποιες φορές το ανάποδο.

  Αναλύοντας τους χρόνους αναμονής και συνδέοντάς τους με τη μεταβλητότητα απαιτείται η εισαγωγή νέων αναλυτικών εργαλείων. Στα επόμενα κεφάλαια θα αναλυθούν κάποια από αυτά. Πιο συγκεκριμένα θα γίνουν προβλέψεις των χρόνων αναμονής, τρόπους μείωσης των χρόνων αναμονής διαλέγοντας τα κατάλληλα επίπεδα χωρητικότητας, ανασχεδιάζοντας το σύστημα εξυπηρέτησης και περιγράφοντας ευκαιρίες να μειωθεί η μεταβλητότητα, όπως επίσης και αποθέματα, ρυθμό ροής και χρόνο ροής.

2.2 Μεταβλητότητα. Από πού έρχεται και πώς μπορεί να μετρηθεί

Αρχικά για την καλύτερη κατανόηση των μέτρων απόδοσης μίας διαδικασίας με την ύπαρξη της πολυπλοκότητας, θα εξεταστεί λίγο περισσότερο η έννοια της μεταβλητότητας. Συγκεκριμένα, θα παρατεθούν οι πηγές μεταβλητότητας και πώς μετριούνται.



Εικόνα 2.1 Μεταβλητότητα και από πού προέρχεται (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Παίρνοντας το παραπάνω διάγραμμα ροής παρατηρείται ότι υπάρχουν τέσσερις πηγές μεταβλητότητας.

1. Η μεταβλητότητα από την εισαγωγή των μονάδων ροής. Η μεγαλύτερη πηγή της μεταβλητότητας σε υπηρεσιακούς οργανισμούς προέρχεται από την ίδια την αγορά. Ενώ κάποια μοτίβα άφιξης πελατών είναι προβλέψιμα (όπως για παράδειγμα η αναχώρηση από ένα ξενοδοχείο πρέπει να γίνεται μέχρι τις 12), πάλι υπάρχει σιγουριά για το πότε θα έρθει ο επόμενος πελάτης.
2. Μεταβλητότητα σε χρόνους επεξεργασίας. Έστω ότι για παράδειγμα υπάρχει σε μία τράπεζα. Όσο προχωράει η ουρά παρατηρούμε ότι κάθε πελάτης χρειάζεται διαφορετικό χρόνο για να εξυπηρετηθεί, ανάλογα με την ανάγκη που έχει. Συμπεραίνεται λοιπόν ότι όταν εμπλέκεται ο ανθρώπινος παράγοντας υπάρχει μεταβλητότητα στους χρόνους επεξεργασίας.
3. Τυχαία μεταβλητότητα των πόρων. Αν οι πόροι υπόκεινται σε τυχαίες κατανομές, για παράδειγμα, αποτυχία μηχανής σε ένα περιβάλλον παραγω-

γής ή απουσία φορέα σε υπηρεσίες παροχής υπηρεσιών, δημιουργείται η μεταβλητότητα.

4. Τυχαία διαδρομή σε περίπτωση πολλαπλών μονάδων ροής της διαδικασίας. Αν για παράδειγμα παρατηρούνται τα επείγοντα ενός νοσοκομείου, μετά την αρχική εξέταση οι ασθενείς οδηγούνται σε διαφορετικές πτέρυγες ανάλογα με τη σοβαρότητα της κατάστασης του καθενός. Ένα πολύ σοβαρό περιστατικό θα οδηγηθεί πιθανόν στην εντατική, ενός κάποιος άλλος με κάτι ελαφρύ θα μπορέσει να εξυπηρετηθεί από έναν γιατρό ή ακόμα και από κάποια νοσοκόμα. Παρότι λοιπόν οι χρόνοι αφίξεων και οι χρόνοι επεξεργασίας είναι ντετερμινιστικοί, το γεγονός ότι υπάρχει τυχαιότητα στην μετέπειτα μεταφορά των ασθενών προκαλεί μεταβλητότητα.

Γενικά, κάθε μορφή μεταβλητότητας μετριέται βάση της τυπικής απόκλισης. Το πρόβλημα βέβαια είναι ότι η τυπική απόκλιση προβάλει ένα *απόλυτο μέτρο* για την μεταβλητότητα. Γι’ αυτό το λόγο είναι πιο κατάλληλο η μέτρηση της μεταβλητότητας να γίνεται με *σχετικούς όρους*. Συγκεκριμένα, ορίζεται ο *συντελεστής μεταβλητότητας* μιας τυχαίας μεταβλητής ως:

$$\text{Συντελεστής τας}=\text{CV}=\frac{\text{Τυπική απόκλιση}}{\text{Μέση τιμή}} \qquad (2.1)$$

Αφού και η τυπική απόκλιση και η μέση τιμή έχουν την ίδια μονάδα μέτρησης, ο συντελεστής της μεταβλητότητας είναι ένα μέγεθος χωρίς μονάδα μέτρησης.

2.3 Διαδικασία αφίξεων

Κάθε διαδικασία απόδοσης που αναλύεται είναι τόσο καλή όσο οι πληροφορίες που την τροφοδοτούνται. Γι’ αυτό είναι πολύ σημαντικό να συλλέγονται και να αναλύονται τα δεδομένα πριν γίνει η χρήση του μαθηματικού μοντέλου. Αρχικά λοιπόν όσο μελετάται η μεταβλητότητα, θα πρέπει να γίνει μία ακριβής αναπαράσταση της ζήτησης, δηλαδή να υπολογιστεί ο χρόνο άφιξης των πελατών.

Πρώτα πρέπει να οριστεί το χρονικό διάστημα κατά το οποίο έχουμε μία άφιξη και το χρόνο άφιξης. Έστω AT_i ο χρόνος άφιξης της ιτης άφιξης. Επίσης, ορίζεται ο χρόνος μεταξύ δύο διαδοχικών αφίξεων (inter-arrival time), IA. Έτσι, $IA_i = AT_{i+1} - AT_i$. Χρησιμοποιώντας τους παραπάνω τύπους συλλέγονται επαρκή δεδομένα για να προχωρήσει η ανάλυσή.

Πριν αναπτυχθεί όμως το μοντέλο που θα οδηγήσει στην πρόβλεψη των αποτελεσμάτων της μεταβλητότητας, πρέπει να γίνει μία ακόμα ανάλυση. Η ανάλυση της διαδικασίας των αφίξεων. Τα ερωτήματα που απασχολούν γι’ αυτή την ανάλυση είναι αν οι διαδικασίες αφίξεων είναι σταθερές, αν ο αναμενόμενος χρόνος άφιξης των πελατών ακολουθεί κάποια κατανομή και αν τα διαστήματα μεταξύ των αφίξεων είναι εκθετικά κατανεμημένα.

Τα παραπάνω ερωτήματα θα απαντηθούν στη συνέχεια όπου ορίζονται οι όροι των σταθερών αφίξεων και των εκθετικά κατανεμημένων διαστημάτων μεταξύ των αφίξεων.

2.3.1 Σταθερές αφίξεις

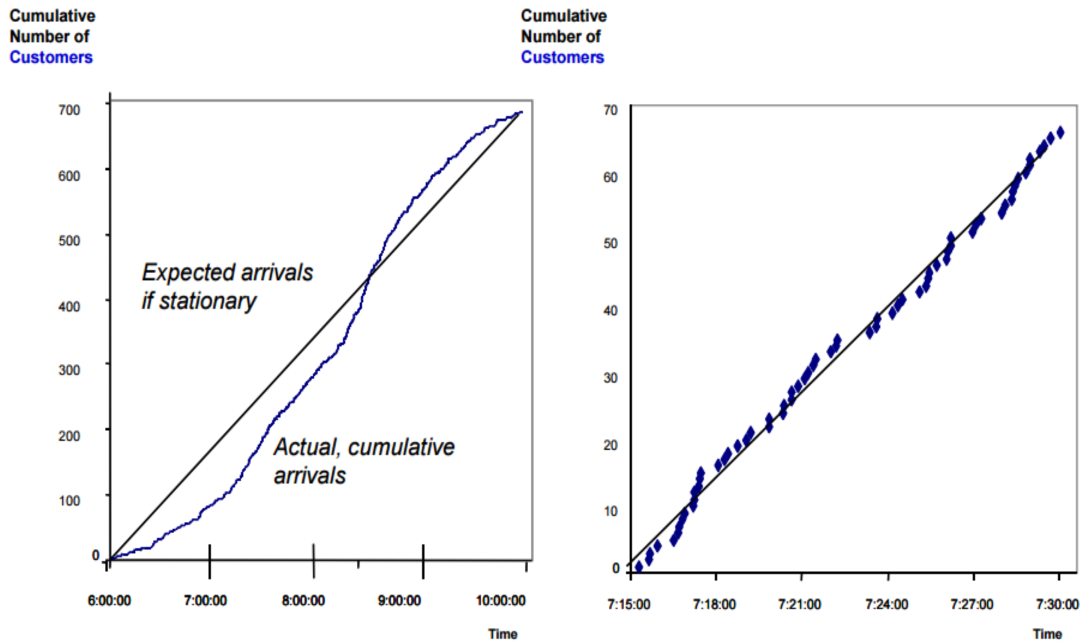
Έστω ότι χρησιμοποιώντας τα παραπάνω συλλέγεται ένας επαρκής αριθμό δεδομένων. Θα ήταν πολύ βολικό να ληφθούν αυτά τα δεδομένα, να εισαχθούν σε ένα υπολογιστικό φύλλο (Excel), να υπολογιστεί η μέση τιμή και η τυπική απόκλιση των διαστημάτων μεταξύ των αφίξεων για μία χρονική περίοδο και να ολοκληρωθεί η ανάλυση απλά και εύκολα. Αν όμως γίνει πιο διεξοδική εξέταση παρατηρείται ότι εκτός από αυτούς τους δύο αριθμούς υπάρχουν και άλλα πράγματα που συμβαίνουν στη διαδικασία και είναι εξίσου σημαντικά. Όπως για παράδειγμα ο αριθμός των αφίξεων μέσα σε ένα χρονικό διάστημα μπορεί να μην είναι σταθερός μέσα στη μέρα.

Γι’ αυτό το λόγο εισάγονται οι παρακάτω ορισμοί:

- Μία διαδικασία άφιξης λέγεται σταθερή όταν δεν αλλάζει σε σχέση με το χρόνο (η κατανομή πιθανότητας που περιγράφει την διαδικασία εισόδου δεν είναι εξαρτημένη από τον χρόνο). Για παράδειγμα μπορεί να μετακινηθεί ένα χρονικό διάστημα συγκεκριμένου μήκους μπρος ή πίσω σε μία χρονική γραμμή χωρίς να αλλάξει ο αναμενόμενος αριθμός αφίξεων.
- Αντίθετα, μία διαδικασία άφιξης η οποία δεν είναι ανεξάρτητη από το χρόνο και παρουσιάζει εποχικότητα ονομάζεται μη σταθερή.

Σε μία διαδικασία άφιξης θα πρέπει να ξεχωρίζονται οι αλλαγές στη ζήτηση που οφείλονται στην εποχικότητα από αυτές που οφείλονται στη μεταβλητότητα και αυτό γιατί η εποχικότητα μπορεί να προβλεφθεί εκ των προτέρων ενώ αυτό δεν ισχύει και για τη μεταβλητότητα. Κάνοντας όμως μια απλή ανάλυση τριών βημάτων φαίνεται αν η διαδικασία είναι σταθερή.

1. Ταξινόμηση όλων των χρόνων αφίξεων έτσι ώστε να αυξάνονται στο χρόνο
2. Σχεδιασμός ενός γραφήματος με $x\ AT_i, y = i$
3. Προσθήκη μίας ευθείας γραμμής από κάτω αριστερά μέχρι πάνω δεξιά



Εικόνα 2.2 Τεστ για σταθερές αφίξεις (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Στην παραπάνω εικόνα υπάρχει ένα παράδειγμα για τις αφίξεις σε ένα τηλεφωνικό κέντρο από τις 6 μέχρι και τις 10 το πρωί. Όπως φαίνεται υπάρχει αρκετά μεγάλη απόκλιση του γραφήματος με την ευθεία γραμμή κάτι που δείχνει ότι η διαδικασία άφιξης σε αυτή την περίπτωση δεν είναι σταθερή αφού παρατηρείται ότι οι κλήσεις που φτάνουν στις 6 το πρωί είναι πολύ λιγότερες απ' ότι στις 10.

Στις περιπτώσεις όπου οι αφίξεις δεν είναι σταθερές καλό είναι τα χρονικά διαστήματα που υπάρχουν να τα διαιρούνται σε μικρότερα. Όταν γίνει αυτό παρατηρείται ότι η εποχικότητα σ΄ αυτά τα μικρότερα διαστήματα είναι πολύ χαμηλότερη. (όπως φαίνεται στο δεξί μέρος της εικόνας 2.2 τα διαστήματα γίναν 15-λεπτα).

Συγκρίνοντας το δεξί με το αριστερό μέρος της εικόνας 2.2 παρατηρείται ότι η διαδικασία άφιξης φαίνεται να είναι σταθερή στα διαστήματα των 15 λεπτών αλλά να παρουσιάζει μεγάλη εποχικότητα κατά τη διάρκεια όλης της ημέρας.

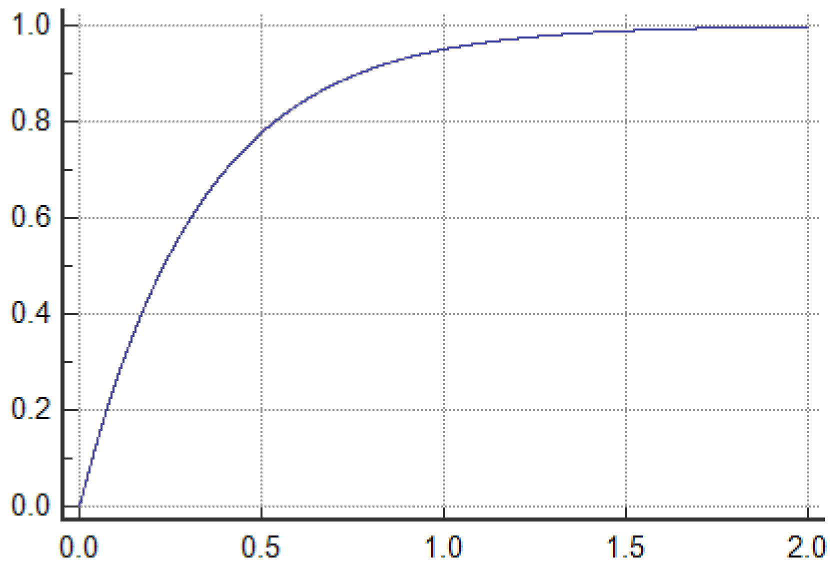
2.3.2 Εκθετικά χρονικά διαστήματα μεταξύ των αφίξεων

Τα χρονικά διαστήματα μεταξύ των αφίξεων συνήθως ακολουθούν την *εκθετική κατανομή*. Αν το IA είναι ένα τυχαίο χρονικό διάστημα μεταξύ των αφίξεων και η διαδικασία μεταξύ των αφίξεων ακολουθεί εκθετική κατανομή έχουμε:

$$\text{Probability } \{IA \leq t\} = 1- e^{-\frac{1}{a}} \qquad (2.2)$$

Όπου α είναι ο μέσος χρόνος μεταξύ των αφίξεων. Οι εκθετικές κατανομές συχνά χρησιμοποιούνται για να μοντελοποιήσουν το χρόνο μεταξύ των αφίξεων στην θεωρία όσο και στην πράξη, και τα δύο τόσο λόγω της καλής προσαρμογής τους με τα εμπειρικά δεδομένα όσο και την ευκολία τους στην ανάλυση. Αν μια διαδικασία άφιξης έχει όντως εκθετικά χρονικά διαστήματα μεταξύ των αφίξεων, αναφερόμαστε σε αυτή σαν την *Διαδικασία Αφίξεων Poisson.*

Μπορεί να φανεί αναλυτικά ότι οι πελάτες φτάνουν στη διαδικασία ανεξάρτητα ο ένας από τον άλλον, δημιουργώντας έτσι ένα μοτίβο ζήτησης με τα διαστήματα μεταξύ των αφίξεων να ακολουθούν την εκθετική κατανομή. Είναι σημαντικό να σημειωθεί πως στην εκθετική κατανομή (που φαίνεται και στην εικόνα 2.3) ο μέσος όρος της είναι ίσος με την τυπική της απόκλιση, α.



Εικόνα 2 3 Εκθετική κατανομή (Πηγή: Cachon- Terwiesch "Matching Supply with Demand")

Μια άλλη σημαντική ιδιότητα της εκθετικής κατανομής είναι γνωστή σαν *ιδιότητα memoryless.* Αυτή η ιδιότητα δείχνει ότι ο αριθμός των αφίξεων στην επόμενη χρονική στιγμή είναι ανεξάρτητος από το πότε συνέβη η τελευταία άφιξη.

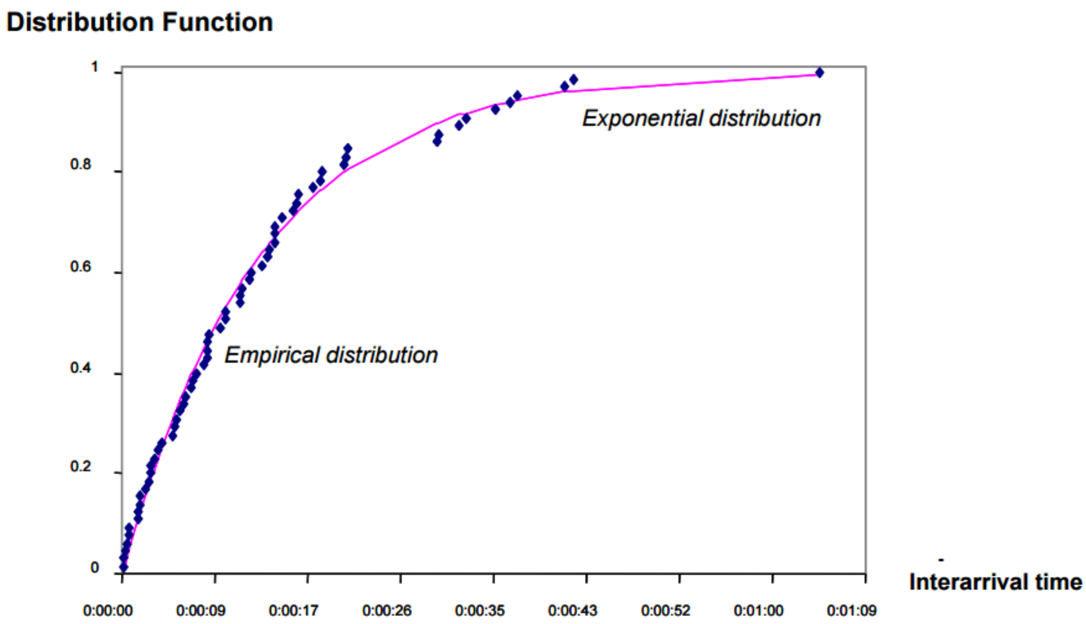
Για να κατανοηθεί καλύτερα αυτή η έννοια, έστω τα επείγοντα ενός νοσοκομείου. Έστω ότι εισέρχεται κατά μέσο όρο ένας ασθενής κάθε 15 λεπτά αλλά τα τελευταία 30 λεπτά δεν έχει εμφανιστεί κανένας. Για μια διαδικασία άφιξης που τα διαστήματα μεταξύ των αφίξεων ακολουθούν την εκθετική κατανομή το γεγονός ότι δεν έχει εμφανιστεί κανείς τα τελευταία 30 λεπτά, δεν επηρεάζει την πιθανότητα του να εμφανιστεί κάποιος μέσα στα επόμενα 15 λεπτά.

Τα παρακάτω τέσσερα βήματα δείχνουν πώς μπορεί να εξεταστεί αν τα διαστήματα μεταξύ των αφίξεων ακολουθούν την εκθετική κατανομή

1. Υπολογίζονται τα χρονικά διαστήματα μεταξύ των αφίξεων $IA_1 \dots \dots IA_n$
2. Ταξινομούνται τα χρονικά διαστήματα μεταξύ των αφίξεων με αύξουσα σειρά. Έστω a_i υποδηλώνει το ιστο μικρότερο χρονικό διάστημα μεταξύ των αφίξεων(a_1 το μικρότερο χρονικό διάστημα μεταξύ των αφίξεων, a_n το μεγαλύτερο)

3. Σχεδιάζονται ζευγάρια ($x = a_i, y = \frac{i}{n}$). Το γράφημα που θα γίνει ονομάζεται συνάρτηση εμπειρικής κατανομής.
4. Συγκρίνεται το γράφημα με μία εκθετική κατανομή με «κατάλληλα επιλεγμένες παραμέτρους». Για να βρεθεί η καλύτερη τιμή για την παράμετρο, τίθεται η παράμετρος της εκθετικής κατανομής ίση με το μέσο χρονικό διάστημα άφιξης που προέκυψε από τα δεδομένα . Αν μερικές παρατηρήσεις από το δείγμα είναι απομακρυσμένες από την προκύπτουσα καμπύλη, θα πρέπει να γίνει προσαρμογή της παραμέτρου της εκθετική κατανομή χειροκίνητα για να ταιριάξει.

Το παρακάτω γράφημα δείχνει τα αποτελέσματα που προκύπτουν από την παραπάνω διαδικασία. Αν η υπογραμμισμένη κατανομή είναι όντως εκθετική, το γράφημα που προκύπτει θα μοιάζει με την αναλυτική κατανομή.



Εικόνα 2.4 Εμπειρική vs. Εκθετική κατανομή (Πηγή: Cachon- Terwiesch "Matching Supply with Demand")

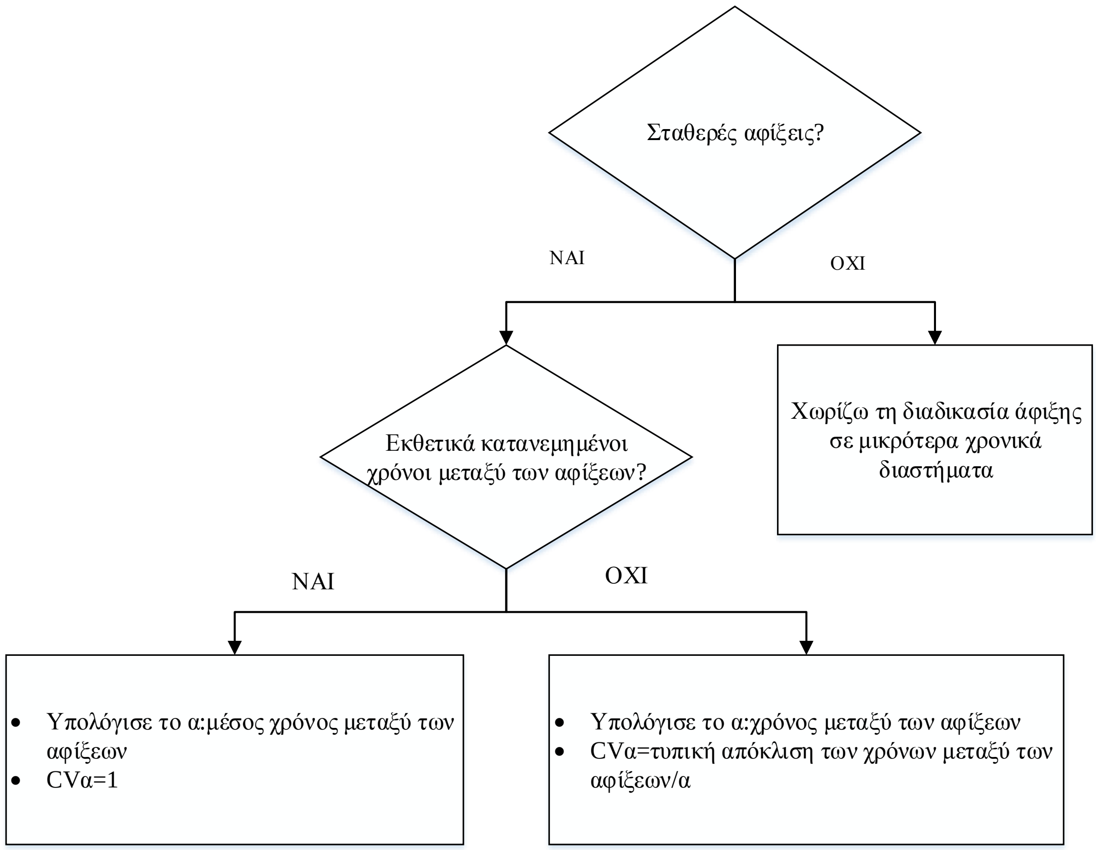
2.3.3 Μη εκθετικά χρονικά διαστήματα μεταξύ των αφίξεων

Σε κάποιες περιπτώσεις οι αφίξεις που πραγματοποιούνται είναι προγραμματισμένες, κάτι που οδηγεί σε μειωμένη μεταβλητότητα, αφού τα διαστήματα μεταξύ των αφίξεων δεν ακολουθούν την κανονική κατανομή. Γι' αυτό ενώ στην περίπτωση της εκθετικής κατανομής η μέση τιμή είναι ίση με την τυπική απόκλιση, εδώ θα πρέπει να υπολογιστούν περισσότερες παράμετροι για να χαρακτηρίσουν την διαδικασία άφιξης.

Ακολουθώντας τον προηγούμενο ορισμό του συντελεστή της μεταβλητότητας, μετριέται η μεταβλητότητα μιας διαδικασίας αφίξεων (ζήτησης) ως:

$$CV_a = \frac{\text{Τυπική απόκλιση χρονικού διασστήματος μεταξύ αφίξεων}}{\text{Μέσο χρονικό διάστημα μεταξύ αφίξεων}} \tag{2.3}$$

Δεδομένου από την εκθετική κατανομή ότι η μέση τιμή είναι ίση με την τυπική από-κλιση, ο συντελεστής της μεταβλητότητας είναι ίσος με 1.



Εικόνα 2.5 Ανάλυση μιας διαδικασίας άφιξης-ζήτησης (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

2.4 Επεξεργασία του χρόνου μεταβλητότητας

Σε μία διαδικασία αφίξεων, εκτός από τη μεταβλητότητα που παρουσιάζεται στην ώρα άφιξης ενός πελάτη, παρατηρείται μεταβλητότητα και στους χρόνους εξυπηρέ-τησης. Αν για παράδειγμα εξεταστούν τα επείγοντα περιστατικά ενός νοσοκομείου, γίνεται αντιληπτό ότι ένας ασθενής μπορεί να χρειαστεί 10 λεπτά για να εξυπηρετη-θεί, ενός κάποιος άλλος να χρειαστεί 30 λεπτά.

Υπήρξαν αναφορές για πολλές διαφορετικές μορφές των κατανομών του χρόνου επεξεργασίας. Εδώ γίνεται χρήση μόνο της μέσης τιμής και της τυπικής τους απόκλισης, και υποτίθεται ότι αυτές οι δύο τιμές περικλείουν όλες τις σχετικές πληροφορίες που χρειαζόμαστε.

Όπως και με τους χρόνους αφίξεων, έτσι και εδώ μπορεί να υπολογιστεί ο συντελεστής μεταβλητότητας ως εξής:

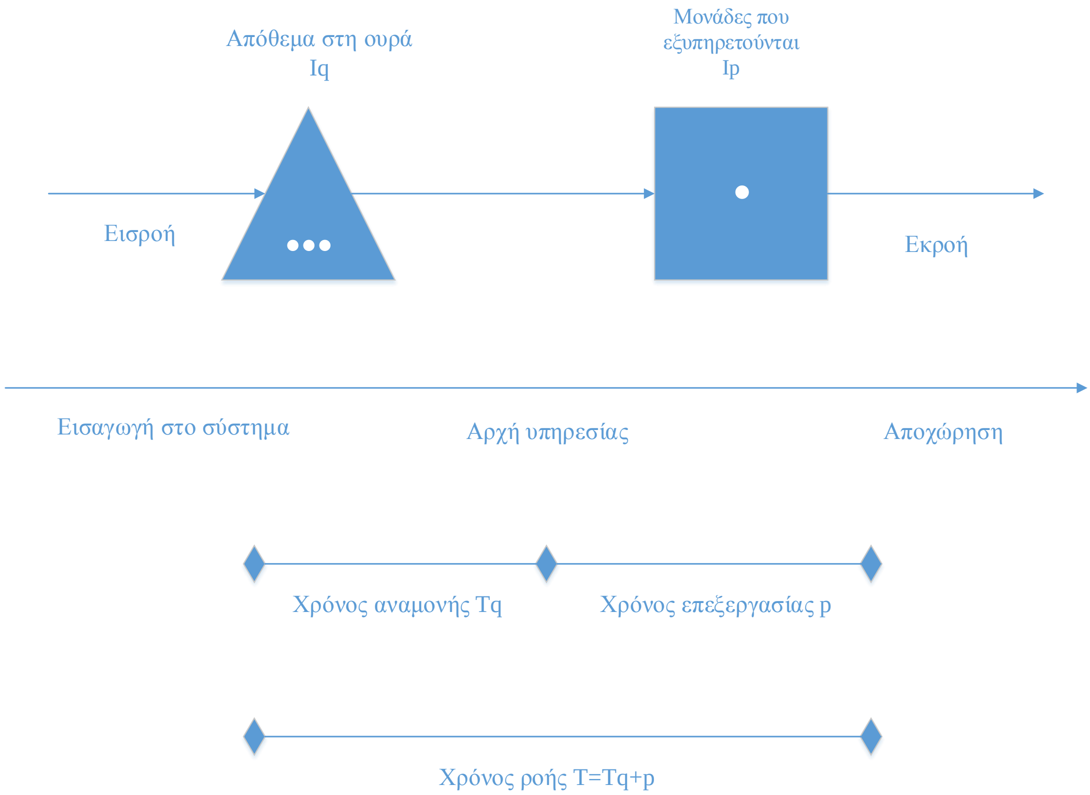
$$CV_p = \frac{Τυπική\ απόκλιση\ του\ χρόνου\ επεξεργασίας}{Μέσος\ χρόνος\ επεξεργασίας} \tag{2.4}$$

Εδώ, ο δείκτης p δηλώνει ότι το CV μετράει την μεταβλητότητα στους χρόνους επεξεργασίας. Όπως και με την διαδικασία άφιξης, πρέπει να μην συγχέεται η μεταβλητότητα με την εποχικότητα.

Τα μοντέλα που εισήχθησαν απαιτούν σταθερή διαδικασία της υπηρεσίας (στην περίπτωση της εποχικότητας στη διαδικασία της υπηρεσίας, χωρίστε τη χρονική γραμμή σε μικρότερα διαστήματα όπως κάναμε στη διαδικασία αφίξεων αλλά δεν απαιτούν άλλες ιδιότητες. Έτσι, η τυπική απόκλιση και η μέση τιμή του χρόνου επεξεργασίας είναι όλα όσα χρειάζεται να είναι γνωστά.

2.5 Μέσος χρόνος αναμονής για την περίπτωση μίας πηγής εξυπηρέτησης

Βάσει των μέτρων μεταβλητότητας που υπάρχουν, εισάγεται ένας απλός τύπος που ανακτά την ικανότητά της πρόβλεψης των βασικών μέτρα απόδοσης της διεργασίας: απόθεμα, ρυθμός ροής και χρόνο ροής.



Εικόνα 2.6 Μία απλή διαδικασία με μία ουρά και έναν εξυπηρετητή (Πηγή: Cachon-Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Στην εικόνα 2.6 φαίνεται το βασικό διάγραμμα διαδικασίας, που αποτελείται από μια ουρά με απεριόριστο χώρο και μία μόνο πηγή εξυπηρέτησης. Οι μονάδες (πελάτες) εισέρχονται στο σύστημα με ένα μοτίβο που παρουσιάζει μεταβλητότητα. Κατά μέσο όρο, μία μονάδα ροής φτάνει κάθε α χρονικές στιγμές. Ορίζεται α ο μέσος χρόνος μεταξύ των αφίξεων. Αυτός ο μέσος όρος αντικατοπτρίζει το μέσο χρόνο μεταξύ των αφίξεων IA_1 μέχρι IA_n. Αφού υπολογιστεί η τυπική απόκλιση των IA_1 μέχρι IA_n χρονικών διαστημάτων μεταξύ των αφίξεων, είναι εύκολο να υπολογιστεί και ο συντελεστής της μεταβολής CV_a της διαδικασίας άφιξης όπως νωρίτερα.

Έστω ότι χρειάζεται περίπου p μονάδες χρόνου για να εξυπηρετηθεί μία μονάδα ροής. Παρόμοια με τη διαδικασία αφίξεων, ορίζονται τα $p_1 \ldots . p_n$ ως οι εμπειρικά παρατηρούμενοι χρόνοι επεξεργασίας και υπολογίζεται ο συντελεστής της μεταβολής για τους χρόνους επεξεργασίας, CV_p. Δεδομένου ότι χρησιμοποιείται μία μόνο πηγή που εξυπηρετεί τις μονάδες ροής που φτάνουν, η παραγωγική ικανότητα του server μπορεί να γραφεί ως 1/P.

Όπως συζητήθηκε στην εισαγωγή, οι περιπτώσεις που αναλύονται είναι αυτές όπου η παραγωγική ικανότητα ξεπερνά το ρυθμό ζήτησης. Έτσι, η προκύπτουσα αξιοποίηση είναι μικρότερη από 100%. Αν η αξιοποίηση ήταν πάνω από 100%, προβλέψιμα θα δημιουργούταν απόθεμα και δεν θα χρειαζόταν κανένα εργαλείο υπολογισμού της μεταβλητότητας για να προβλεφτεί ότι οι μονάδες ροής θα επιβαρύνουν τους χρόνους αναμονής. Ωστόσο, το πιο σημαντικό όμως είναι ότι οι μονάδες ροής επιβαρύνουν το χρόνο αναμονής ακόμα και αν ο εξυπηρετητής χρησιμοποιείται κάτω από το 100%.

Δεδομένου ότι η παραγωγική ικανότητα ξεπερνά τη ζήτηση και υποθέτοντας ότι ποτέ δεν χάνουμε πελάτη, υπάρχει περιορισμός ως προς τη ζήτηση και έτσι ο ρυθμός ροής R είναι ο ρυθμός ζήτησης. Συγκεκριμένα, αφού φτάσει ο πελάτης, κατά μέσο όρο, κάθε α μονάδες χρόνου, ο ρυθμός ροής R=1/α.

Σε αυτό το σημείο γίνεται εισαγωγή της έννοιας της αξιοποίησης, η οποία είναι βασική για την ανάλυσή . Αξιοποίηση είναι ένα μέτρο που δείχνει πόσο παράγει η διαδικασία σε σχέση με το πόσο θα μπορούσε να παράγει αν λειτουργούσε στο μέγιστο δυνατό και υπολογίζεται ως εξής:

$$\text{Αξιοποίηση}=\frac{\text{Ρυθμός ροής}}{\text{Χωρητικότητα}}=\frac{1/\alpha}{1/p}=\frac{p}{a}<100\% \tag{2.5}$$

Έστω τώρα η περίπτωση μιας μονάδας η οποία κινείται σε ένα σύστημα (όπως στην εικόνα 2.6). Έστω T_q ο χρόνος που περνάει η μονάδα περιμένοντας να εξυπηρετηθεί. Ο δείκτης q υποδηλώνει ότι αυτό είναι μόνο ο χρόνος που περνάει η μονάδα περιμένοντας στην ουρά. Έτσι, ο T_q δεν περιλαμβάνει τον πραγματικό χρόνο επεξεργασίας, ο οποίος ορίστηκε p. Βάσει του χρόνου αναμονής στην ουρά T_q και του μέσου χρόνου επεξεργασίας p, υπολογίζεται ο χρόνος ροής (το χρόνο που θα περάσει η μονάδα ροής μέσα στο σύστημα) ως

$$\begin{gathered}\text{Flow time= Χρόνος στην ουρά+ Χρόνος επεξεργασίας}\\ \text{T}=T_q+\text{p}\end{gathered} \tag{2.6}$$

Αντί για την προοπτική της μονάδας ροής, μπορεί επίσης το σύστημα να εξετασθεί σαν ολότητα, και να τεθεί το ερώτημα πόσες μονάδες ροής θα είναι στην ουρά και πόσες θα εξυπηρετούνται. Έστω I_q το απόθεμα που είναι στην ουρά και I_p ο αριθμός των μονάδων που εξυπηρετούνται. Αφού το απόθεμα I_q στην ουρά και το απόθεαμα I_p στη διαδικασία είναι τα μόνα μέρη που μπορεί να υπάρχει απόθεμα, υπολογίζεται το συνολικό απόθεμα στο σύστημα I=I_q+I_p.

Το I_p στην περίπτωση της μίας πηγής που εξετάζεται μπορεί να είναι μόνο 0 ή 1. Αυτό συμβαίνει γιατί είτε θα υπάρχει μία μονάδα που εξυπηρετεί ($I_p = 1$) είτε όχι ($I_p = 0$). Η πιθανότητα μια τυχαία χρονική στιγμή ο εξυπηρετητής να είναι απασχολημένος, ανταποκρίνεται στην αξιοποίηση. Για παράδειγμα, αν η αξιοποίηση της διαδικασίας είναι 40%, υπάρχει πιθανότητα 0,4 σε μία τυχαία στιγμή ο εξυπηρετητής να είναι απασχολημένος. Εναλλακτικά έστω ότι από τα 60 λεπτά της ώρας ο εξυπηρετητής είναι απασχολημένος για

0,4*60[λεπτά/ ώρα]=24 λεπτά

Ενώ είναι φανερό ότι το απόθεμα που εξυπηρετείται I_p και ο χρόνος επεξεργασίας p είναι εύκολο να υπολογιστούν, δεν ισχύει το ίδιο και για το απόθεμα στην ουρά I_q και το χρόνο αναμονής στην ουρά T_q.

Όντας γνωστά ο χρόνος επεξεργασίας p,η αξιοποίηση και η μεταβλητότητα, όπως μετρήθηκαν με τους συντελεστές διακύμανσης CV_a και CV_p, μπορούν να υπολογιστούν ο μέσος χρόνος αναμονής στην ουρά με τη χρήση του παρακάτω τύπου:

Χρόνος στην ουρά(T_q)= $Χρόνος\ επεξεργασίας * \left(\frac{Αξιοποίηση}{1-Αξιοποίηση}\right) * \left(\frac{{CV_a}^2+{CV_p}^2}{2}\right)$ (2.7)

Ο τύπος δεν υποδεικνύει ότι οι χρόνοι επεξεργασίας ή τα χρονικά διαστήματα μεταξύ των αφίξεων ακολουθούν μία συγκεκριμένη κατανομή. Όμως, για την περίπτωση των μη εκθετικών χρόνων μεταξύ των αφίξεων, ο τύπος μόνο προσεγγίζει τον αναμενόμενο χρόνο στη ουρά, σε αντίθεση με την 100% ακρίβεια. Ο τύπος θα έπρεπε να χρησιμοποιείται μόνο για τις περιπτώσεις των σταθερών διαδικασιών.

Η παραπάνω εξίσωση υποστηρίζει πως ο χρόνος αναμονής στην ουρά είναι το γινόμενο τριών συντελεστών:

- Του χρόνου αναμονής ο οποίος εκφράζεται σαν πολλαπλάσιο του χρόνου επεξεργασίας. Ωστόσο, είναι σημαντικό να έχουμε στο μυαλό μας ότι ο χρόνος επεξεργασίας επηρεάζει άμεσα την αξιοποίηση.
- Της επίδρασης της αξιοποίησης. . Να σημειωθεί ότι η αξιοποίηση πρέπει να είναι λιγότερο από 100%. Αν είναι ίση ή μεγαλύτερη από 100%, η ουρά συνεχίζει να μεγαλώνει. Αυτό δεν καθοδηγείται από την μεταβλητότητα, αλλά από το ότι δεν έχει την απαιτούμενη παραγωγική ικανότητα. Παρατηρείται ότι ο συντελεστής της αξιοποίησης δεν είναι γραμμικός και μεγαλώνει όσο το επίπεδο αξιοποίησης πλησιάζει το 100%. Για παράδειγμα, για αξιοποίησης =0,5 ο συντελεστής αξιοποίησης είναι 0,5/(1-0,5)=1. Για αξιοποίηση=0,6, ο συντελεστής είναι 0,6/(1-0,6)=1,5 και για αξιοποίηση=0,65, ο συντελεστής ανεβαίνει στο 0,65/(1-0,65)=1,85.
- Του ποσού της μεταβλητότητας στο σύστημα που μετριέται από το μέσο όρο του τετραγώνου του συντελεστή απόκλισης των χρόνων μεταξύ των αφίξεων CV_a και των χρόνων εξυπηρέτησης CV_p. Αφού τα CV_a και CV_p δεν επηρεάζουν ούτε το μέσο χρόνο εξυπηρέτησης p ούτε τη αξιοποίηση u, φαίνεται ότι ο χρόνος αναμονής αυξάνεται με τη μεταβλητότητα στο σύστημα.

Αφού έχει βρεθεί το T_q παρακάτω παρατίθεται ένας ακόμα τύπος υπολογισμού του αποθέματος (I)

$$I = R * T = \frac{1}{a} * (T_q + p) \tag{2.8}$$

Για ορθότερα όμως αποτελέσματα πρέπει να δειχθεί πως υπολογίζονται και τα I_q και I_p.

Ο υπολογισμός του I_q γίνεται αν παρατηρηθεί η γραμμή αναμονής σαν μία μικρή διαδικασία και γνωρίζοντας το T_q :

$$I_q = \frac{1}{a} * T_q \tag{2.9}$$

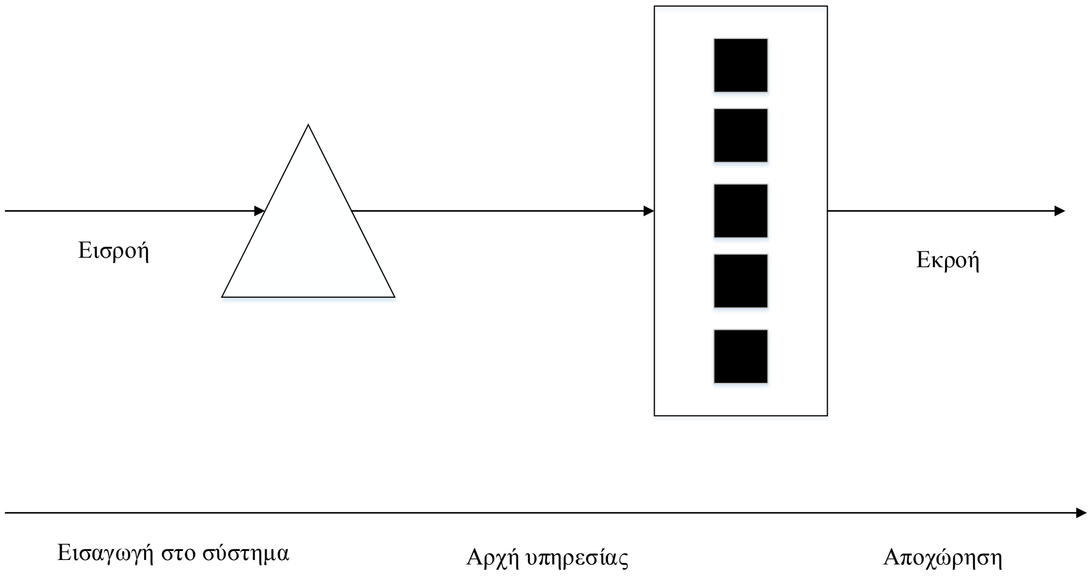
Τέλος, θα πρέπει να παρατηρηθεί ότι αφού υπάρχει μία μόνο πηγή εξυπηρέτησης, σε κάθε χρονική στιγμή θα εξυπηρετείται και μία μόνο μονάδα. Υπάρχουν όμως και περιπτώσεις που δεν εξυπηρετείται καμία μονάδα, καθώς η αξιοποίηση της πηγής είναι πολύ πιο κάτω από 100%. Έτσι λοιπόν ο μέσος όρος των μονάδων που εξυπηρετούνται υπολογίζεται ως εξής:

$$I_p = \text{Probability}\,\{0\ \mu ονάδες\ εξυπηρετούνται\} * 0 + Probability\,\{1\ \mu ονάδα\ εξυπηρετείται\} * 1 \tag{2.10}$$

$$I_p = (1-u) * 0 + u * 1 = u$$

2.6 Μέσος χρόνος αναμονής για την περίπτωση πολλών πηγής εξυπηρέτησης

Έχοντας αναλύσει το χρόνο αναμονής με την παρουσία της μεταβλητότητας για μία πολύ απλή διαδικασία, με μόνο μία ουρά και μία πηγή, τώρα αναλύονται πιο πολύπλοκα συστήματα. Συγκεκριμένα, αναπτύσσεται ένα μοντέλο χρόνου αναμονής μιας διαδικασίας που αποτελείται από μία περιοχή αναμονής (ουρά) και ένα στάδιο της διαδικασίας που εκτελείται από πολλές, ίδιες πηγές.



Εικόνα 2.7 Διαδικασία με μία ουρά και πολλούς παράλληλους εξυπηρετητές (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Στην παραπάνω εικόνα φαίνεται η διάταξη αυτής της διαδικασίας. Όπως και στην περίπτωση με μία πηγή, χρειάζονται οι περιπτώσεις όπου η αξιοποίηση είναι κάτω από 100%.

Έστω m ο αριθμός των παράλληλων πηγών (εξυπηρετητών). Δεδομένου αυτού, αντιμετωπίζεται μία κατάσταση όπου ο μέσος χρόνος εξυπηρέτησης είναι πιθανό να είναι πολύ μεγαλύτερος απ' ότι ο μέσος χρόνος μεταξύ των αφίξεων. Οι m πηγές έχουν παραγωγική ικανότητα m/p, καθώς ο ρυθμός ζήτησης δίνεται πάλι από τον τύπο 1/α. Μπορεί να υπολογιστεί η αξιοποίηση u ως

$$Αξιοποίηση = \frac{Ρυθμός\ Ροης}{Χωρητικότητα} = \frac{\frac{1}{Χρόνος\ μετξύ\ αφίξεων}}{\frac{Αριθμός\ πηγών}{Χρόνος\ εξυπηρέτησης}} = \frac{\frac{1}{a}}{\frac{m}{p}} = \frac{p}{a * m} \qquad (2.11)$$

Η μονάδα ροής σταδιακά θα ξοδέψει T_q μονάδες χρόνου αναμονής για εξυπηρέτηση. Μετά μετακινείται στην επόμενη διαθέσιμη πηγή, όπου ξοδεύει p μονάδες χρόνου εξυπηρέτησης. Όπως πριν, ο συνολικός χρόνος ροής είναι το άθροισμα του χρόνου αναμονής και του χρόνου εξυπηρέτησης:

Χρόνος ροής= Χρόνος αναμονής στην ουρά+ Χρόνος εξυπηρέτησης

$$T = T_q + p$$

Βάσει του χρόνου εξυπηρέτησης p, της αξιοποίησης u, των συντελεστών απόκλισης για την εξυπηρέτηση (CV_p) και της διαδικασίας άφιξης (CV_α) όπως και του αριθμού των πηγών στο σύστημα (m), μπορεί να υπολογιστεί ο μέσος χρόνος αναμονής T_q με τη χρήση του ακόλουθου τύπο:

$$\begin{aligned} & \text{Χρόνος στην ουρά} \\ & = \left(\frac{\text{Χρόνος εξυπηρέτησης}}{m}\right) \\ & * \left(\frac{\text{Αξιοποίηση}^{\sqrt{2(m+1)}-1}}{1-\text{Αξιοποίηση}}\right) * \left(\frac{{CV_a}^{2}+{CV_p}^{2}}{2}\right) \end{aligned} \tag{2.12}$$

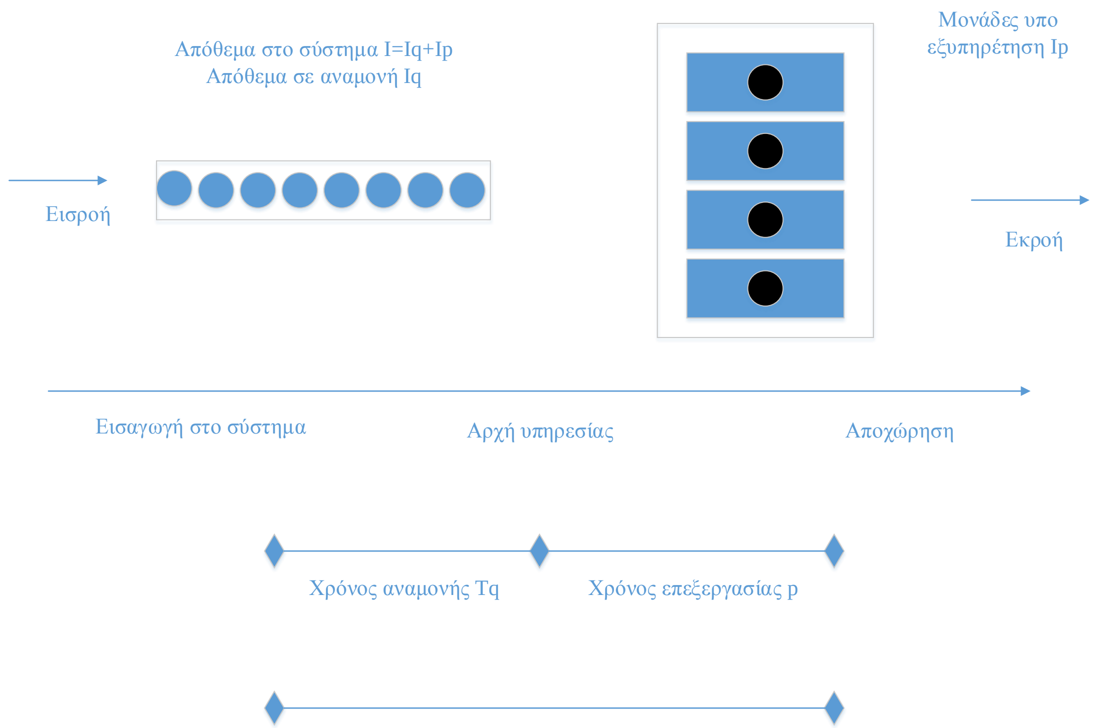
Ενώ η παραπάνω εξίσωση δεν είναι και τόσο εύκολη στη χρήση της, μπορεί εύκολα να αποτυπωθεί σε ένα υπολογιστικό φύλλο. Επίσης, υπολογίζει το μέσο χρόνο αναμονής για το σύστημα που αλλιώς θα χρειαζόταν πολύ πιο εξειδικευμένα προγράμματα για να υπολογιστεί.

Όπως στην περίπτωση με τη μία πηγή, ο χρόνος αναμονής εκφράζεται ως το γινόμενο του χρόνου επεξεργασίας, ενός συντελεστή αξιοποίησης και ενός συντελεστή μεταβλητότητας. Επίσης παρατηρείται ότι για την ειδική περίπτωση όπου m=1, ο παραπάνω τύπος είναι ακριβώς ίδιος με τον τύπο του χρόνου αναμονής για μία πηγή. Να σημειωθεί ότι όλα τα άλλα μέτρα απόδοσης, συμπεριλαμβανομένου και του χρόνου ροής (T), του αποθέματος στο σύστημα (I) και του αποθέματος στην ουρά(I_q), μπορούν να υπολογιστούν όπως είδαμε πριν.

Στην περίπτωση της μίας πηγής η εξίσωση του χρόνου αναμονής μας δίνει μία ακριβή ποσοτικοποίηση του χρόνου αναμονής καθώς ο χρόνος μεταξύ των αφίξεων ακολουθεί εκθετική κατανομή. Αντίθετα όμως, στην περίπτωση με τους πολλούς εξυπηρετητές, ο χρόνος αναμονής είναι μόνο μία προσέγγιση. Η εξίσωση δουλεύει για τα περισσότερα στοιχεία που υπολογίζουμε, και ειδικότερα αν η αναλογία της αξιοποίησης u με τον αριθμό των εξυπηρετητών m είναι μεγάλη. (u/m είναι μεγάλο)

Τώρα που υπολογίστηκε ο χρόνος αναμονής, μπορεί να υπολογιστεί ο μέσος αριθμός μονάδων ροής στη χρονική περίοδο I_q, ο μέσος αριθμός μονάδων ροής στην υπηρεσία I_p και ο μέσος αριθμός μονάδων ροής σε ολόκληρο το σύστημα I=I_q+I_p.

Παρακάτω είναι μία εικόνα με τα βασικά μέτρα απόδοσης και θα γίνει και μία σύνοψη αυτών.



Εικόνα 2.8 Σύνοψη των μέτρων απόδοσης (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

1. Συλλέγονται τα παρακάτω δεδομένα

- Αριθμός εξυπηρετητών m
- Χρόνος εξυπηρέτησης, p
- Χρόνος μεταξύ των αφίξεων, α
- Συντελεστής απόκλισης για χρόνο μεταξύ των αφίξεων, CV_a, και για τον χρόνο εξυπηρέτησης, CV_p

2. Υπολογίζεται η αξιοποίησης: $u = \frac{p}{a*m}$
3. Υπολογίζεται ο αναμενόμενος χρόνο αναμονής

$$T_q = \left(\frac{Χρόνος\ επεξεργασίας}{m}\right) * \left(\frac{Αξιοποίηση^{\sqrt{2(m+1)}-1}}{1 - Αξιοποίηση}\right) * \left(\frac{{CV_a}^2 + {CV_p}^2}{2}\right)$$

4. Βάσει του T_q, υπολογίζονται τα μέτρα απόδοσης που έμειναν
   Χρόνος ροής(flow time) T= T_q+p
   Απόθεμα στη υπηρεσία(inventory in service) $I_p = m * u$
   Απόθεμα στην ουρά (inventory in queue) $I_q = T_q/a$
   Απόθεμα στο σύστημα (inventory in the system) I= $I_p + I_q$

Εφόσον οι πηγές είναι περισσότερες από μία, οι μονάδες ροής μπορεί να είναι και περισσότερες από μία. Αν u είναι η αξιοποίηση της διαδικασίας, είναι επίσης και η αξιοποίηση κάθε μίας από τις m πηγές. Ο αριθμός των μονάδων ροής για κάθε μία από τις m πηγές υπολογίζεται και με τον παρακάτω τύπο:

u*1+(1-u)*0=u

Προσθέτοντάς το στις m πηγές προκύπτει

Απόθεμα σε διαδικασία= Αριθμός πηγών * Αξιοποίηση

I_p=m*u

2.7 Επίπεδα εξυπηρέτησης στα προβλήματα του χρόνου αναμονής

Ενώ μέχρι τώρα το βασικό θέμα ενασχόλησης είναι κυρίως ο χρόνος αναμονής, το πιο σημαντικό πρακτικά δεν είναι πόσο χρόνο θα περιμένει ο πελάτης στην ουρά και πόσο χρόνο θα κάνει να εξυπηρετηθεί, αλλά το πώς θα βιώσει ο ίδιο ο άνθρωπος αυτή την αναμονή.

Για παράδειγμα, στο τηλεφωνικό κέντρο μίας τράπεζας κάποιος μπορεί να περιμένει μέχρι και 20 λεπτά για να εξυπηρετηθεί. Ένα τέτοιο γεγονός σίγουρα θα τον δυσαρεστήσει και μπορεί μέχρι και να τον κάνει να κλείσει το τηλέφωνο, ακόμα και αν υπάρχει ηχογραφημένο μήνυμα που να λέει πόσος περίπου είναι ο χρόνος αναμονής.

Γι΄ αυτό το λόγο δεν θα πρέπει να υπολογιστεί μόνο ο μέσος χρόνος αναμονής, αλλά και την πιθανότητα αυτός ο χρόνος να ξεπεράσει κάποιο στόχο. (target wait time-TWT). Ορίζεται λοιπόν το επίπεδο εξυπηρέτησης για έναν δοσμένο TWT ως την εκατοστιαία αναλογία των πελατών που θα εξυπηρετηθούν στον TWT ή μικρότερες μονάδες του χρόνου αναμονής:

Επίπεδο εξυπηρέτησης= Probability {Χρόνος αναμονής ≤TWT } (2.13)

Αυτό το μέτρο εξυπηρέτησης βοηθάει να φανεί κατά πόσο μία υπηρεσία είναι σε θέση να ανταποκριθεί στη ζήτηση σε ένα συγκεκριμένο χρονικό όριο. Ένα επίπεδο εξυπηρέτησης στο 90 τοις εκατό για έναν TWT=30 δευτερόλεπτα σημαίνει ότι αυτό το 90 τοις εκατό των πελατών εξυπηρετούνται σε λιγότερα από 30 δευτερόλεπτα του χρόνου αναμονής.

Τα παραπάνω επίπεδα εξυπηρέτησης είναι ένα συνηθισμένο μέτρο απόδοσης για τις επιχειρήσεις παροχής υπηρεσιών. Χρησιμοποιούνται πολύ συχνά από εταιρίες που θέλουν να αναθέσουν αλλού την υπηρεσία (outsource) σαν ένα τρόπο να εντοπίζουν την ανταπόκριση του φορέα παροχής των υπηρεσιών τους.

Δεν υπάρχει κάποιος κανόνας για το πιο επίπεδο εξυπηρέτησης είναι σωστό για μία επιχείρηση παροχής υπηρεσιών. Για παράδειγμα, ως απάντηση για την μεγάλη πίεση του κοινού, το σιδηροδρομικό σύστημα της Γερμανίας, υιοθέτησε μία πολιτική που λέει ότι το 30 τοις εκατό των κλήσεων στον αριθμό παραπόνων θα πρέπει να εξυπηρετείται μέσα σε 20 δευτερόλεπτα. Η ταχύτητα απάντησης μίας εταιρίας στις κλήσεις εξαρτάται από τη θέση αλλά και το πόσο σημαντικές είναι οι κλήσεις για την εταιρία. Το σιδηροδρομικό σύστημα της Γερμανίας, ας πούμε, μπορεί για άλλες εταιρίες να είναι απαράδεκτο.

2.8 Οικονομικές επιπτώσεις : Δημιουργώντας ένα πρόγραμμα προσωπικού

Μέχρι τώρα έχει γίνει η ανάλυση μίας διαδικασίας εξυπηρέτησης για ένα δεδομένο αριθμό εκπροσώπων εξυπηρέτησης πελατών και παρατέθηκαν οι τρόποι για να προβλεφθούν οι επικείμενες ώρες αναμονής. Αυτό γεννά την ερώτηση πόσοι εξυπηρετητές θα πρέπει να υπάρχουν ώστε να είναι ικανοποιημένες οι μονάδες ροής (πελάτες). Όσοι περισσότεροι εξυπηρετητές υπάρχουν τόσο λιγότερη θα είναι η ώρα αναμονής αλλά τόσο περισσότερο θα πρέπει οι πληρωμές να γίνουν σε ωρομίσθια.

Για να βρεθεί η λύση θα πρέπει να υπολογιστούν τέσσερα πράγματα:

- Το κόστος αναμονής που θα έχει η υπηρεσία
- Το κόστος υπηρεσιών που προκύπτει από τον αριθμό των διαθέσιμων εξυπηρετητών
- Το κόστος για εκείνους τους πελάτες που για κάποιο λόγο δεν θα μπορέσουν ούτε και να μπουν στην ουρά
- Το κόστος των πελατών που έχουν κουραστεί στην αναμονή και εγκαταλείπουν το σύστημα

Για να αποφασιστεί πόσοι εξυπηρετητές θα υπάρχουν, πρέπει πρώτα να εξεταστεί πόσο εξυπηρετική θέλει να είναι η εταιρία με τους πελάτες της. Αρχικά λοιπόν πρέπει να εξετάσει πόσο μέσο χρόνο αναμονής θέλει περίπου να έχει η εταιρία. Στη συνέχεια θα πρέπει για ένα δεδομένο αριθμό αφίξεων να εξετάσει πόσους εξυπηρετητές θα πάρει ώστε να επιτύχει το μέσο χρόνο αναμονής που έθεσε.

Η επίτευξη καλύτερης εξυπηρέτησης γίνεται με τη θυσία της αυξημένης εργασίας. Όσο περισσότεροι εξυπηρετητές προγραμματίζονται να εξυπηρετήσουν τόσο μικρότερη είναι η χρήση τους.

Υπολογίζεται τώρα το κόστος άμεσης εργασίας για το οποίο πρέπει να είναι γνωστά δύο πράγματα:

1. Ο αριθμός των μονάδων ροής που εισέρχονται σε κάθε μονάδα του χρόνου(flow rate)
2. Το ποσό του μισθού που πληρώνουμε για την ίδια χρονική περίοδο

Έτσι :

$$\text{Κόστος άμεσης εργασίας} = \frac{\text{ΣΥΝΟΛΙΚΑ ΩΡΟΜΙΣΘΙΑ ΑΝΑ ΜΟΝΑΔΑ ΕΡΓΑΣΙΑΣ}}{\text{ΡΥΘΜΟΣ ΡΟΗΣ ΑΝΑ ΜΟΝΑΔΑ ΕΡΓΑΣΙΑΣ}} \quad (2.14)$$

Όπου τα συνολικά ωρομίσθια ανά μονάδα εργασίας είναι ο αριθμός των εξυπηρετητών ,m, επί το ρυθμό των ωρομισθίων τους (το ποσό που εισπράττουν σε κάποια μονάδα του χρόνου π. χ ώρες, λεπτά) και ο ρυθμός ροής έχει σχέση το ρυθμό άφιξης (1/α).

Ένας εναλλακτικός τρόπος γραψίματος του κόστους άμεσης εργασίας χρησιμοποιεί τον ορισμό της αξιοποίησης. (u=p/(a × m)).

Αυτός ο τρόπος παρουσίασης του κόστους άμεσης εργασίας έχει μία διαισθητική ερμηνεία: Ο πραγματικός χρόνος εξυπηρέτησης, p, αυξάνεται κατά ένα συντελεστή 1/αξιοποίηση για να εξηγήσει σωστά το χρόνο αδράνειας. Για παράδειγμα εάν η αξιοποίηση είναι 50 % η χρέωση είναι με 1 % ποινή χρόνου αδράνειας για κάθε ένα 1$ που ξοδεύεται σε παραγωγική εργασία. Στην περίπτωση εδώ η χρήση είναι 66% άρα το κόστος άμεσης εργασίας είναι:

$$Κόστος\ άμεσης\ εργασίας\ = \frac{1{,}5\frac{λεπτά}{κλήση} \times 16.66\frac{cents}{λεπτό}}{0.66} = 38 \text{ cents/κλήση}$$

Από αυτόν τον υπολογισμό βρίσκονται οι συνέπειες εξόδων που έχουν να κάνουν με διάφορα σενάρια που αφορούν το προσωπικό. Θα πρέπει άρα να δημιουργηθεί ένα κατάλληλο πλάνο προσωπικού ώστε να επιτευχθεί όσο το δυνατόν χαμηλότερο κόστος άμεσης εργασίας.

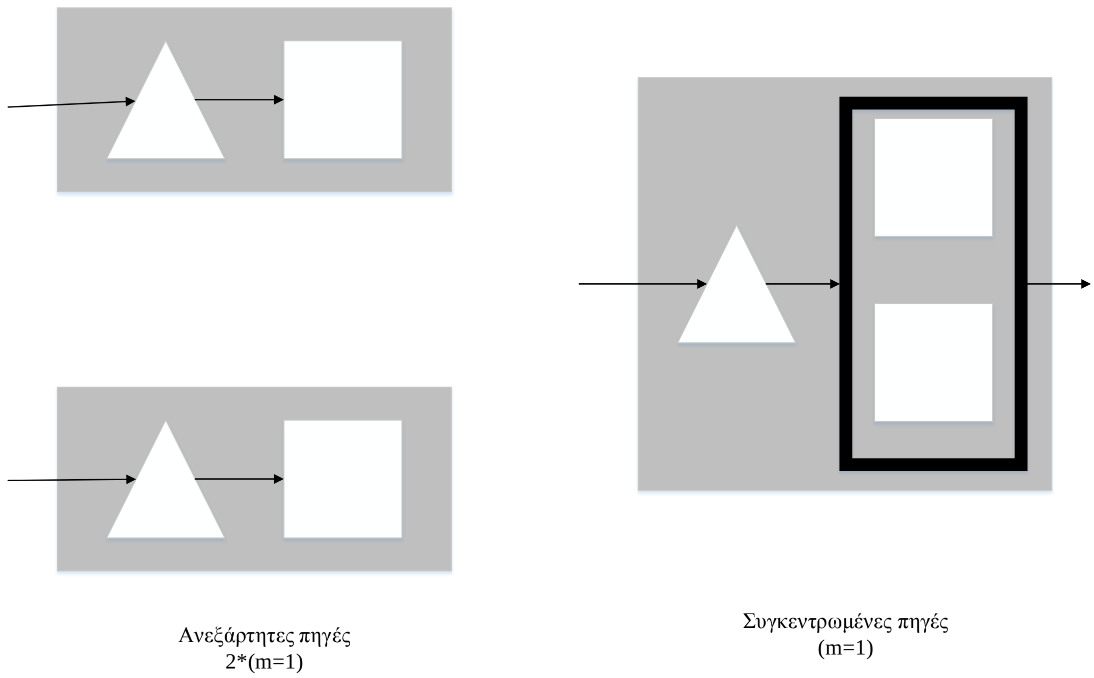
Θα πρέπει όμως πρώτα να εξεταστεί και η περίπτωση με μετακινούμενη διαδικασία άφιξης. Σε αυτήν την περίπτωση ένα κοινό πρόβλημα που υπάρχει είναι ότι δεν είναι γνωστό σε πόσα διαστήματα θα πρέπει να χωριστεί η γραμμή του χρόνου ώστε να επιτευχθεί μία σταθερή διαδικασία άφιξης. Σε αυτήν την περίπτωση το συμπέρασμα είναι ότι θα πρέπει τα διαστήματα να είναι αρκετά μεγάλα έτσι ώστε να υπάρχουν αρκετά στοιχεία ώστε να γίνουν αξιόπιστες εκτιμήσεις για το ρυθμό άφιξης του διαστήματος και να υπάρχει αρκετός χρόνος ώστε η ουρά να φτάσει σε μία σταθερή κατάσταση.

Άρα, στην πράξη η εύρεση ενός πλάνου διαχείρισης προσωπικού μπορεί να είναι πιο περίπλοκη αφού πρέπει να δίνει λογαριασμό για:

- Διαλείμματα χειρίστων
- Διάρκεια περιόδου εργασίας. Συνήθως δεν είναι δυνατόν να ζητηθεί από ένα χειριστή να έρθει για δουλεία μόνο για μια ώρα. Είτε κάποιος θα πρέπει να παρέχει περισσότερες ώρες είτε να διοχετεύσει προσωρινά τις κλήσεις σε άλλα μέλη της οργάνωσης (επιτηρητής, υπάλληλοι υποστήριξης)

2.9 Επιπτώσεις συγκέντρωσης: Οικονομίες κλίμακας

Έστω μία διαδικασία που ανταποκρίνεται σε δύο (m) διαδικασίες άφιξης της ζήτησης που επεξεργάζονται από δύο (m) πανομοιότυπους εξυπηρετητές. Αν δεν μπορεί η ζήτηση να εξυπηρετηθεί άμεσα, τότε περιμένει μπροστά από τον εξυπηρετητή όπου έφτασε αρχικά. Ένα τέτοιο σύστημα φαίνεται στην παρακάτω εικόνα (αριστερά).



Εικόνα 2.9 Η έννοια της συγκέντρωσης (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Αν τώρα αυτά τα δύο συστήματα συνδυαστούν σε ένα με μία ουρά αναμονής και δύο πανομοιότυπους εξυπηρετητές (Εικόνα 2.9 δεξιά) τότε εμφανίζεται η έννοια της συγκέντρωσης (pooling).

Σε αυτή την περίπτωση για τον υπολογισμό του χρόνου αναμονής γίνεται χρήση του αντίστοιχου τύπο για πολλούς εξυπηρετητές

$$T_q = \left(\frac{Χρόνος\ επεξεργασίας}{m}\right) * \left(\frac{Αξιοποίηση^{\sqrt{2(m+1)}-1}}{1 - Αξιοποίηση}\right) * \left(\frac{{CV_a}^2 + {CV_p}^2}{2}\right)$$

Με αντίστοιχο υπολογισμό της αξιοποίησης

$$u = \frac{p}{a \times m}$$

Παρακάτω είναι ένα αριθμητικό παράδειγμα για τη διαφορά των δύο συστημάτων που παρουσιάζει η εικόνα 2.9.

Έστω δύο μικρές υπηρεσίες τροφίμων. Θεωρείται ότι και οι δυο τους έχουν μια ροή άφιξης του πελάτη, με μέσο χρόνο μεταξύ των αφίξεων α= 5 λεπτά και ένα συντελεστή απόκλισης που ισούται με ένα. Ο χρόνος επεξεργασίας p είναι 4 λεπτά ανά πελάτη και ο συντελεστής μεταβολής για τη διαδικασία είναι επίσης ίσος με ένα. Κατά συνέπεια, οι δύο υπηρεσίες τροφίμων αντιμετωπίζουν αξιοποίηση των p / α = 0,8 σύμφωνα με τους τύπους για την περίπτωση της μίας πηγής εξυπηρέτησης (αφού η καθεμία δρα ανεξάρτητα από την άλλη). Και επίσης για το χρόνο αναμονής είναι :

$$T_q = Χρόνος\ επεξεργασίας * \left(\frac{Αξιοποίηση}{1-Αξιοποίηση}\right) * \left(\frac{{CV_a}^2+{CV_p}^2}{2}\right) = 4 * \frac{0,8}{1-0,8} * \frac{1+1}{2} = 16\ λεπτά$$

Έστω τώρα το ίδιο παράδειγμα αν συνδυαστούν τις δύο υπηρεσίες. Η ικανότητα της ομαδοποιημένης διαδικασίας έχει αυξηθεί κατά έναν παράγοντα του δύο και τώρα είναι 2/3 της μονάδας ανά λεπτό. Ωστόσο, το ποσοστό της ζήτησης, επίσης, έχει διπλασιαστεί. Αν υπήρχε ένας πελάτης κάθε πέντε λεπτά που φθάνει για την υπηρεσία 1 και έναν πελάτη κάθε πέντε λεπτά που φθάνει για την υπηρεσία 2, η συγκεντρωτική υπηρεσία βιώνει ένα ρυθμό άφιξης ενός πελάτη κάθε 2,5 λεπτά. Η αξιοποίηση της ομαδοποιημένης διαδικασίας υπολογίζεται ως εξής:

$$u = \frac{p}{a \times m}$$

$$= 6/\ (2.5 \times\ 2) = 0.8$$

Παρατηρείται ότι και στην περίπτωση της ομαδοποίησης η αξιοποίηση είναι ίδια. Ο συνδυασμός των δύο διαδικασιών, με αξιοποίηση του 80 τοις εκατό οδηγεί σε ομαδοποιημένο σύστημα, με ποσοστό αξιοποίησης 80 τοις εκατό. Αν υπολογιστεί όμως ο χρόνος αναμονής θα έχουμε:

$$T_q = \left(\frac{Χρόνος\ επεξεργασίας}{m}\right) * \left(\frac{Αξιοποίηση^{\sqrt{2(m+1)}-1}}{1-Αξιοποίηση}\right) * \left(\frac{{CV_a}^2+{CV_p}^2}{2}\right)$$
$$= \frac{6}{2} * \frac{0,8^{\sqrt{2(2+1)}-1}}{1-0,8} * \frac{1+1}{2} = 10,85\ λεπτά$$

Άρα φαίνεται ότι με τη συγκέντρωση εξυπηρετείται ο ίδιος αριθμός πελατών, στον ίδιο χρόνο επεξεργασίας αλλά με πολύ μικρότερο χρόνο αναμονής.

Παρόλα τα οφέλη που φαίνονται όμως παραπάνω, η συγκέντρωση δεν πρέπει να θεωρείται ως κάτι το ιδανικό. Σε κάποιες περιπτώσεις τα οφέλη της είναι πολύ χαμηλότερα όπως για παράδειγμα αν σε μία τράπεζα δουλεύουν δύο υπάλληλοι, ένας στις απλές συναλλαγές και ένας στις πιο δύσκολες περιπτώσεις, και τοποθετηθούν μαζί αυτό μπορεί να είναι επιζήμιο χρονικά γιατί μπορεί κάποια στιγμή να τύχει μια δύσκολη περίπτωση σε αυτόν που είναι για τις απλές και να μην ξέρει να την χειριστεί και έτσι να χαθεί πολύτιμος χρόνος. Το ίδιο θα μπορούσε να ισχύει και σε ένα δικηγορικό γραφείο με πολλούς δικηγόρους. Άρα η συγκέντρωση δεν είναι για όλες τις περιπτώσεις.

2.10 Μείωση της μεταβλητότητας

Μέχρι τώρα έχει αποδειχθεί ότι η μεταβλητότητα είναι ο εχθρός όλων των διαδικασιών (κανένα μέτρο απόδοσης δεν βελτιώνεται όσο αυξάνεται η μεταβλητότητα). Έτσι αντί να γίνεται απλά δεκτή η μεταβλητότητα, πρέπει να βρεθούν οι τρόποι μείωσής της.

2.10.1 Τρόποι μείωσης της μεταβλητότητας των αφίξεων

Ένας τρόπος μείωσης της μεταβλητότητας είναι η εφαρμογή των συστημάτων ραντεβού (appointment systems , αναφέρονται επίσης και ως συστήματα κρατήσεων σε κάποιους κλάδους). Αυτός είναι ένας τρόπος ταύτισης της προσφοράς με τη ζήτηση. Με τα συστήματα των ραντεβού μπορεί να μειωθεί η μεταβλητότητα αφού οι πελάτες έρχονται σε συγκεκριμένες χρονικές στιγμές. Παρόλα αυτά, δεν πρέπει να παραβλέπονται και κάποια προβλήματα που υπάρχουν.

Ένα από τα προβήματα είναι ότι τα ραντεβού δεν μειώνουν πάντα τη μεταβλητότητα των αφίξεων. Αυτό συμβαίνει γιατί κάποιος πελάτης μπορεί να μην εμφανιστεί στην ώρα ή ακόμα να μην εμφανιστεί και καθόλου. Θα πρέπει λοιπόν κάθε σύστημα κρατήσεων που σέβεται τον εαυτό του, να μπορεί να διαχειριστεί τέτοιες καταστάσεις είτε με κάποια χρηματική ποινή είτε με μεγαλύτερο χρόνο αναμονής γι’ αυτούς που αργούνε. Βέβαια αυτό έρχρται σε σύγκρουση με το τι είναι δίκαιο και αποδεκτό. Για παράδειγμα σε ένα ιατρείο αν μπορεί να αργήσει ο γιατρός γιατί να μην μπορεί και ο ασθενής.

Ένα άλλο πρόβλημα είναι ότι δεν είναι γνωστό τι ποσοστό της χωρητικότητας θα πρέπει να διατεθεί στα ραντεβού. Αυτό γιατί πάντα θα υπάρχουν και έκτακτες περιπτώσεις. Όπως για παράδειγμα σε ένα νοσοκομείο, τα επείγοντα περιστατικά δεν έχουν κλείσει κάποιο ραντεβού και σίγουρα προηγούνται των ασθενών που έχουν κάνει κάποια κράτηση.

Ο μεγαλύτερος όμως προβληματισμός είναι ότι ενώ το σύστημα κρατήσεων φαίνεται να μειώνει τη μεταβλητότητα της διαδικασίας αφίξεων, δεν μειώνει τη μεταβλητότητα της πραγματικής ζήτησης. Για παράδειγμα ένας γιατρός του ΕΟΠΥΥ ο οποίο λειτουργεί με ραντεβού, φαινομενικά μειώνει τη μεταβλητότητα. Πρακτικά όμως κάτι τέτοιο δεν συμβαίνει γιατί ο ασθενής που έχει σήμερα ραντεβού μπορεί να το έχει κλείσει πριν δύο μήνες. Έτσι φαίνεται ότι με τα στυστήματα κρατήσεων κρύβεται μία αναντιστοιχία μεταξύ προσφοράς και ζήτησης γι' αυτό πριν εφαρμόσουμε κάποιο τέτοιο σύστημα θα πρέπει να μετράμε δύο πράγματα:

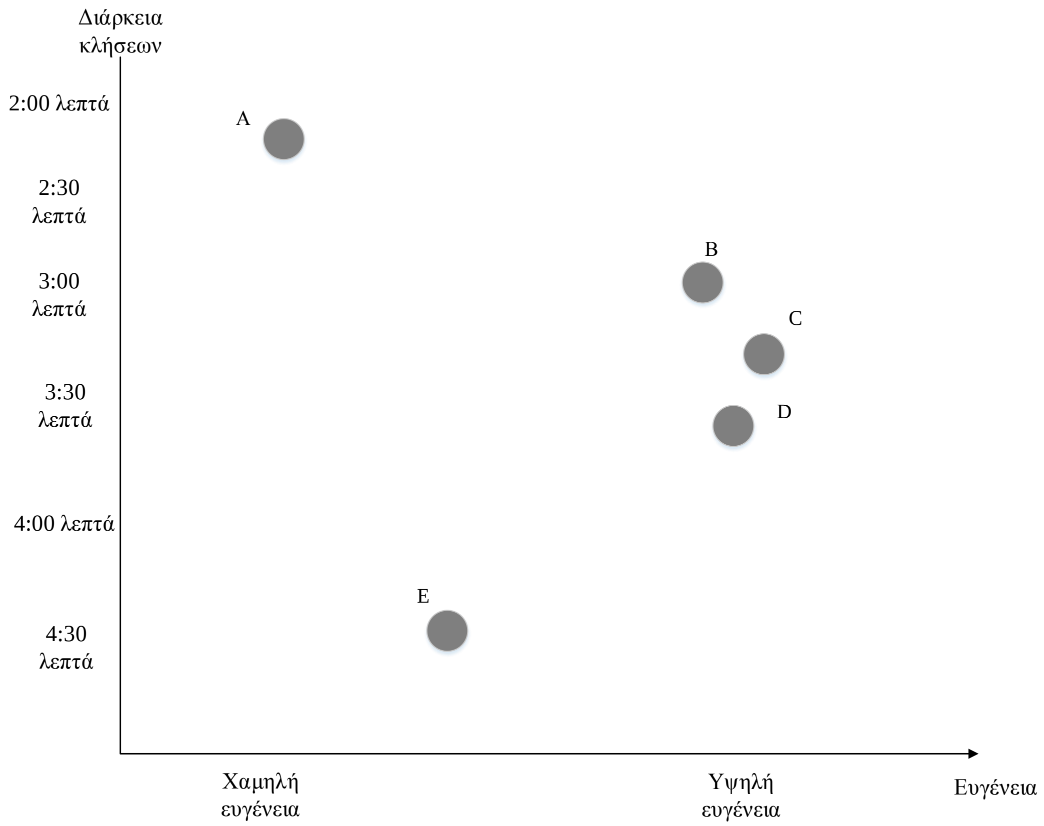
1. Τον αριθμό τον πελατών που έχουν κλείσει ραντεβού και τώρα περιμένουν εκείνη τη μέρα για να επισκεφτούν το γιατρό
2. Τον αριθμό των πελατών που περιμένουν στον προθάλαμο του ιατρείου για να κλείσουν κάποιο ραντεβού.

Εκτός από τα συστήματα κρατήσεων, μπορεί να μειωθεί η μεταβλητότητα στο χρόνο άφιξης των πελατών και με κάποια κίνητρα ώστε να αποφεύγουν τις ώρες αιχμής. Παραδείγματα κάποιων τέτοιων κινήτρων είναι: εκπτώσεις σε δωμάτια ξενοδοχείων σε περιόδους που δεν είναι υψηλής ζήτησης, μειωμένα αεροπορικά εισητήρια σε μέρες που δεν ταξιδεύει πολύς κόσμος (κυρίως καθημερινές) και πολλά άλλα.

2.10.2 Τρόποι μείωσης της μεταβλητότητας του χρόνου

Εκτός από τη μείωση της μεταβλητότητας από την αλλαγή συμπεριφοράς των πελατών, πρέπει να ελεγθεί και πως θα μειωθεί η εσωτερική μεταβλητότητα. Ωστόσο, αν γίνει προσπάθεια σταθεροποίησης των δραστηριοτήτων (μείωση του συντελεστή διακύμανσης των χρόνων επεξεργασίας) ή μείωσης των χρόνων επεξεργασίας, πρέπει να ισορροπηθούν η λειτουργική αποτελεσματικότητα και η ποιότητα υπηρεσιών που βιώνει ο πελάτης.

Υποθετικά τώρα ότι υπάρχει μία συγκεκριμένη υπηρεσία κλήσεων με πέντε έστω τηλεφωνητές, για μία τηλεφωνική εταιρία. Έστω οι τηλεφωνητές A, B, C, D, E.



Εικόνα 2.10 Επίδοση των τηλεφωνητών ανάλογα με τη διάρκεια της κλήσης και τη συμπεριφορά (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Στην παραπάνω εικόνα φαίνεται ότι οι B, C, D έχουν σχετικά σύντομης διάρκειας κλήσεις αλλά είναι φιλικοί προς τους πελάτες. Ο τηλεφωνητής Α έχει σύντομης διάρκειας κλήσεις αλλά δεν βαθμολογείται καλά σε σχέση με την ευγένεια. Τέλος ο φορέας Ε, παρουσιάζει κλήσεις μεγάλης διάρκειας αλλά βαθμολογείται πολύ μέτρια για την ευγένειά του προς τους πελάτες.

Από όλα τα παραπάνω παρατηρείται ότι όσο μεγαλύτερη είναι η διάρκεια των κλήσεων, τόσο λιγότερη ευγένεια παρατηρούμε από τη μεριά των τηλεφωνητών. Επίσης υπάρχουν φορείς που δεν είναι ούτε γρήγοροι ούτε και ευγενικοί.

Για να μειωθούν αυτά τα προβλήματα πολλοί οργανισμοί χρησιμοποιούν συστήματα εκπαίδευσης και τεχνολογίας. Για παράδειγμα στα περισσότερα τηλεφωνικά κέντρα έχουν συστήματα εκπαίδευσης στα οποία δίνουν στους φορείς του κάποια κείμενα για να τα μάθουν και να μπορούν να απαντήσουν με ευγένεια σε οτιδήποτε τους πει κάποιος πελάτης.

Υπάρχουν όμως και άλλες ευκαιρίες για βελτίωση που προσανατολίζονται κυρίως προς τη μείωση της μεταβλητότητας των χρόνων επεξεργασίας:

- Παρά το γεγονός ότι σε ένα περιβάλλον υπηρεσιών ο φορέας πρέπει να αναγνωρίσει την ιδιοσυγκρασία του κάθε πελάτη, ο φορέας μπορεί ακόμα να ακολουθήσει μια σταθερή διαδικασία. Για παράδειγμα, ένας ταξιδιωτικός πράκτορας σε ένα τηλεφωνικό κέντρο θα μπορούσε να χρησιμοποιήσει προκαθορισμένο κείμενο (σενάρια) για την αλληλεπίδραση του με τον πελάτη (καλωσόρισμα, πρώτη ερώτηση, δυναμικό κλείσιμο στο τέλος της συνομιλίας. Έτσι, το να είσαι γνώστης σχετικά με τη διαδικασία (πότε να πεις τι) είναι εξίσου σημαντικό όσο οι γνώσεις σχετικά με το προϊόν.
- Οι χρόνοι επεξεργασίας σε ένα υπηρεσιακό περιβάλλον σε αντίθεση με τους χρόνους επεξεργασίας σε ένα κατασκευαστικό πλαίσιο, δεν είναι υπό τον πλήρη έλεγχο του πόρου. Ο πελάτης παίζει ένα σημαντικό ρόλο στην δραστηριότητα του πόρου ο οποίος εισάγει αυτόματα ένα ορισμένο ποσό της μεταβλητότητας .Ποια είναι η συνέπεια; Τουλάχιστον από την άποψη της μεταβλητότητας, η απάντηση είναι σαφής: Μειώστε τη συμμετοχή του πελάτη κατά τη διάρκεια της υπηρεσίας σε ένα σπάνιο πόρο, όπου είναι δυνατόν.
- Τέλος, πολλές φορές υπάρχει μεταβλητότητα του χρόνου επεξεργασίας στην ποιότητα. Σε περιβάλλοντα παραγωγής, αυτό θα μπορούσε να περιλαμβάνει επεξεργασία εκ νέου μιας μονάδας που αρχικά δεν πληροί τις προδιαγραφές. Ωστόσο, αυτό επίσης συμβαίνει σε οργανισμούς παροχής υπηρεσιών (π.χ., ένας ασθενής ο οποίος απελευθερώνεται από τη μονάδα εντατικής θεραπείας, αλλά αργότερα η εισαγωγή του ξανά στην εντατική φροντίδα μπορεί να θεωρηθεί ως επεξεργασία εκ νέου).

Κεφάλαιο 3: Μεταβλητότητα στην απόδοση της διαδικασίας: Απώλειες απόδοσης

Το παρόν κεφάλαιο προέρχεται κυρίως από το βιβλίο Matching Supply with Demand των Cachon- Terwiesch και το κεφάλαιο 9. Αρχικά απαντάται το γιατί δεν λειτουργούν οι μέσοι όροι. Στη συνέχεια αναφέρεται η ανυπομονησία τον μονάδων ροής (πελατών) και η απόδοση της απώλειας. Τέλος, παρουσιάζονται κάποιες πηγές με σειριακή μεταβλητότητα.

3.1 Εισαγωγή

Εκτός από αυτά που αναλύθηκαν πριν, η μεταβλητότητα δημιουργεί ένα ακόμα πρόβλημα: τις απώλειες απόδοσης. Στις περιπτώσεις όπου προκύπτουν οι απώλειες απόδοσης, η ζήτηση είναι μεγαλύτερη από την χωρητικότητα. Κάθε μονάδα ροής που φτάνει, δεν μπορεί ή δεν θέλει να περιμένει με αποτέλεσμα να εγκαταλείπει το σύστημα αν δεν εξυπηρετηθεί αμέσως. Για παράδειγμα, σε ένα κομμωτήριο αν όλες οι κομμώτριες είναι απασχολημένες ο πελάτης που θα έρθει και θα το δει θα φύγει και θα πάει σε κάποιο άλλο κομμωτήριο.

Εξετάζεται η πιο απλή περίπτωση όπου δεν υπάρχει καθόλου χώρος αναμονής, αφού οι περιπτώσεις με απώλειες απόδοσης είναι πιο δύσκολο να εξεταστούν σε σχέση με αυτές που οι πελάτες θέλουν να περιμένουν. Παρακάτω φαίνονται κάποια αναλυτικά εργαλεία και μερικά παραδείγματα.

3.2 Γιατί οι μέσοι όροι δεν λειτουργούν

Παρακάτω παρατίθεται τώρα ένα παράδειγμα. Έστω μία καφετέρια που πουλάει μόνο καφέδες στο χέρι. Η ζήτηση κατά μέσο όρο είναι ένας καφές κάθε πέντε λεπτά. Η ζήτηση όμως δεν μπορεί να είναι σταθερή, γιατί κάποια στιγμή μπορεί να φτάσουν 2 πελάτες μαζί ενώ κάποια άλλη στιγμή να μην έρθει κανείς. Όπως και να’ χει όμως οι πελάτες σίγουρα δεν θέλουν να περιμένουν.

Πίνακας 3.1 Παράδειγμα μεταβλητότητας (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Σενάριο	**Ζήτηση**	**Προσφορά**	**Ρυθμός Ροής**
A	0	0	0
B	0	1	0
C	0	2	0
D	1	0	0
E	1	1	1
F	1	2	1
G	2	0	0
H	2	1	1
I	2	2	2
Μέσος Όρος	1	1	$\frac{5}{9}$

Η δυναμικότητα που οδηγεί στην προσφορά, μπορεί επίσης να πάρει τιμές 0,1 και 2 αλλά ο μέσος όρος της είναι και πάλι 1.

Από μία συγκεντρωτική οπτική γωνία η ζήτηση και η προσφορά φαίνεται να είναι ίσες και κατά μέσο όρο ο πωλητής θα έπρεπε να πουλάει περίπου 1 καφέ τα 5 λεπτά.

$$\text{Ρυθμός ροής} = \text{Minimum } \{\text{Ζήτηση, Προσφορά}\} = \text{Minimum } \{1,1\} = 1 \tag{3.1}$$

Στον πίνακα 3.1 υπολογίζεται το ρυθμό ροής και μετά τους μέσους όρους προσφοράς και ζήτησης και έτσι προέκυψαν 9 περιπτώσεις. Φαίνεται ότι στις τρεις πρώτες περιπτώσεις δεν πουλιέται κανένας καφές ενώ στις τελευταίες τρεις η ζήτηση υπερβαίνει την προσφορά. Έτσι ενώ η ζήτηση αυξάνεται, πουλιέται περίπου ένας καφέ κάθε πέντε λεπτά.

Παρατηρώντας το μέσο ρυθμό ροής που υπολογίστηκε με αυτόν τον τρόπο, παρατηρείται ότι κοντά στις μισές πωλήσεις που ήταν αναμενόμενο να γίνουν σύμφωνα με τη συγκεντρωτική ανάλυση, δεν πραγματοποιήθηκαν! Ο λόγος είναι ο εξής: για να πραγματοποιηθεί μία πώληση χρειάζονται δύο παράγοντες την ίδια στιγμή: η ζήτηση (ένας πελάτης) και η προσφορά (η ικανότητα του πωλητή να φτιάξει τον καφέ). Αν όμως ο πωλητής είχε φτιάξει κάποιους καφέδες όταν δεν είχε κόσμο και τους είχε αποθηκεύσει τότε ο ρυθμός ροής θα ήταν πολύ καλύτερος. Έτσι λοιπόν με αυτό το παράδειγμα αποδείχθηκε ότι τελικά οι μέσοι όροι δεν λειτουργούν.

3.3 Απώλειες απόδοσης για μία απλή διαδικασία

Έστω ένα σύστημα με περιορισμένο αριθμό πηγών (εξυπηρετητών) . Σε αυτή την περίπτωσή αν το σύστημα είναι γεμάτο τότε οι μονάδες ροής (πελάτες) παραπέμπονται σε άλλο σύστημα που έχει ελεύθερο χώρο. Για παράδειγμα σε ένα νοσοκομείο έστω ότι υπάρχουν τέσσερις πτέρυγες στα επείγοντα. Αν είναι όλες γεμάτες οι ασθενείς που φτάνουν δεν μπορούν να περιμένουν γιατί πολλές φορές ακόμα και πέντε λεπτά αναμονής μπορεί να αποβούν μοιραία. Σε αυτήν την περίπτωση λοιπόν, το νοσοκομείο θα πρέπει να δηλώσει ότι βρίσκεται σε κατάσταση πληρότητας και έτσι να σταλούν οι ασθενείς σε κάποιο άλλο κοντινό νοσοκομείο όπου υπάρχει επαρκής χωρητικότητα.

Έστω λοιπόν τώρα ότι στο σύστημά εισέρχεται πελάτης κάθε α μονάδες χρόνου και χρειάζεται p μονάδες χρόνου για να εξέλθει από το σύστημα. Δεδομένου ότι όλες οι πηγές είναι ελεύθερες η χωρητικότητα του συστήματος είναι:

$$Χωρητικότητα = \frac{Αριθμός\ πηγών}{Χρόνος\ επεξεργασίας} \tag{3.2}$$

Οι μονάδες ροής φτάνουν με τυχαία σειρά κι έτσι χρησιμοποιούνται εκθετικοί χρόνοι μεταξύ των αφίξεων και κατά συνέπεια συναντάται ένας συντελεστής μεταβλητότητας $CV_\alpha = 1$.

Ένα πάρα πολύ σημαντικό βήμα για την ανάλυσή είναι ο υπολογισμός της πιθανότητας να υπάρχουν στο σύστημα m μονάδες ροής, P_m. Η πιθανότητα είναι μεγάλης σημασίας, καθώς μόλις βρεθούν m μονάδες στο σύστημα, το σύστημα θα πρέπει να παραπέμψει αλλού οποιοδήποτε εισερχόμενο αίτημα μέχρι να τελειώσει με κάποια μονάδα ροής. Η πιθανότητα να είναι και οι m εξυπηρετητές απασχολημένοι, P_m εξαρτάται από δύο μεταβλητές:

- Την έμμεση αξιοποίηση(implied utilization). Δεδομένου ότι κάποιοι πελάτες δεν γίνονται δεκτοί στην διαδικασία, δεν χρειάζεται να εξεταστεί η περίπτωση όπου ο δείκτης προσφοράς ξεπερνά τον δείκτη της ζήτησης (1/α). Αυτή η υπόθεση ήταν απαραίτητη στο προηγούμενο κεφάλαιο, γιατί αλλιώς η γραμμή αναμονής θα είχε «εκραγεί». Στο σύστημα όπου αυτόματα τερματίζει την διαδικασία σε περίπτωση υψηλής ζήτησης, αυτό δεν συμβαίνει. Ως εκ τούτου, το u τώρα περιλαμβάνει την περίπτωση όπου η αξιοποίηση(utilization) υπερβαίνει το 100%, και γι΄ αυτό γίνεται αναφορά για έμμεση αξιοποίηση(implied utilization) (Δείκτης ζήτησης/ Χωρητικότητα) σε αντίθεση με την αξιοποίηση (Ρυθμός ροής/ Χωρητικότητα).

- Τον αριθμό των πηγών m.Ξεκινώντας την ανάλυσή υπολογίζεται η έμμεση αξιοποίηση(implied utilization):

$$u = \frac{Ρυθμός\ ζήτησης}{Χωρητικότητα} \tag{3.3}$$

Βάσει της έμμεσης αξιοποίησης (implied utilization) u και του αριθμού των πηγών m, χρησιμοποιείται η ακόλουθη μέθοδος για τον υπολογισμό της πιθανότητας να είναι απασχολημένοι όλοι οι m εξυπηρετητές, P_m. Ορίζουμε r=u*m=p/α.

Πίνακας απωλειών Erlang							
m							
		1	**2**	**3**	**4**	**5**	**6**
	0.10	0.0909	0.0045	0.0002	0.0000	0.0000	0.0000
	0.20	0.1667	0.0164	0.0011	0.0001	0.0000	0.0000
	0.25	0.2000	0.0244	0.0020	0.0001	0.0000	0.0000
	0.30	0.2308	0.0335	0.0033	0.0003	0.0000	0.0000
r	0.33	0.2500	0.0400	0.0044	0.0004	0.0000	0.0000
	0.40	0.2857	0.0541	0.0072	0.0007	0.0000	0.0000
	0.50	0.3333	0.0769	0.0127	0.0016	0.0002	0.0000
	0.60	0.3750	0.1011	0.0198	0.0030	0.0004	0.0000
	0.67	0.4000	0.1176	0.0255	0.0042	0.0006	0.0001
	0.70	0.4118	0.1260	0.0286	0.0050	0.0007	0.0001
	0.75	0.4286	0.1385	0.0335	0.0062	0.0009	0.0001

Με τη βοήθεια του πίνακα του τύπου απώλειας του Erlang θα βρεθεί η πιθανότητα να είναι όλες οι πηγές m απασχολημένες και ως εκ τούτου μια καινούρια μονάδα ροής που θα φτάσει θα πρέπει να απορριφθεί. Πρώτα, βρίσκεται η αντίστοιχη γραμμή κεφαλίδας (r) που δείχνει την αναλογία του χρόνου επεξεργασίας προς τα διαστήματα μεταξύ των αφίξεων (βλέπε πίνακα 2). Ύστερα, βρίσκεται η αντίστοιχη κεφαλίδα στήλης (m) που δείχνει τον αριθμό των πηγών. Η τομή αυτής της γραμμής με αυτήν την στήλη είναι

$$Probability\{all\ m\ servers\ busy\} = P_m(r)$$

Για το παράδειγμα του νοσοκομείου που αναφέρθηκε στην αρχή, αυτός ο τύπος δείχνει το ποσοστό του χρόνου που το νοσοκομείο θα βρίσκεται σε κατάσταση πληρότητας.

Καταγράφονται κάποιες παρατηρήσεις για να εξηγηθεί το αντίκτυπο του δείκτη του χρόνου επεξεργασίας προς το χρόνο μεταξύ των αφίξεων, r, και του αριθμού των πηγών m στην πιθανότητα όλοι οι εξυπηρετητές να είναι απασχολημένοι:

- Η πιθανότητα $P_m(r)$ και επομένως και η ανάλυση δεν απαιτούν τον συντελεστή διακύμανσης για τη διαδικασία της υπηρεσίας. Η ανάλυση ισχύει για την (ρεαλιστική) περίπτωση των εκθετικά κατανεμημένων χρόνων μεταξύ των αφίξεων. Παρόλα αυτά, γίνεται η υπόθεση ότι ο συντελεστής διακύμανσης για τη διαδικασία αφίξεων ισούται με ένα.
- Ο τύπος που στηρίζεται στον πίνακα του παραρτήματος αποδίδεται στον Agner Krarup Erlang, έναν Δανό μηχανικό που εφηύρε πολλά από τα μοντέλα που χρησιμοποιούμε στα προηγούμενα κεφάλαια για τον εργοδότη του, το Τηλεφωνικό κέντρο της Κοπεγχάγης. Σε αυτήν την περίπτωση, οι αφίξεις ήταν εισερχόμενες κλήσεις για τις οποίες

υπήρχε είτε μια τηλεφωνική γραμμή διαθέσιμη είτε όχι(σε αυτή την περίπτωση οι κλήσεις χάνοντας, γι΄ αυτό και ο τύπος ονομάστηκε τύπος απώλειας του Erlang).

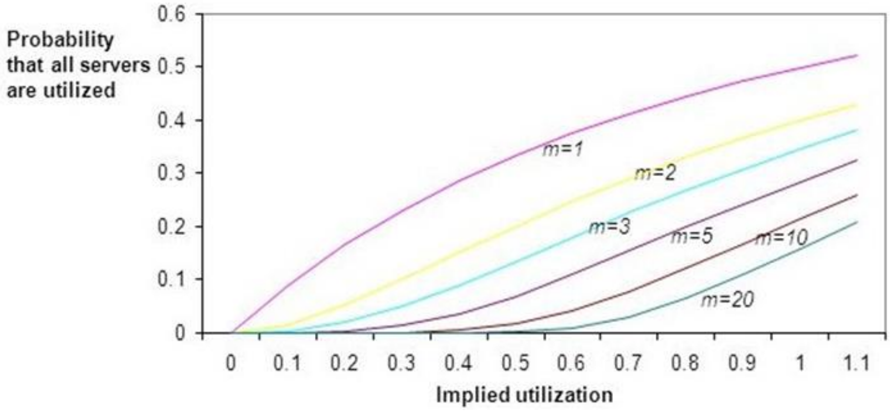
- Στο πρώτο κεφάλαιο, παρατέθηκε ο τύπος που οδηγεί στον πίνακα του τύπου απώλειας του Erlang. Μπορεί να γίνει χρήση του τύπου κατευθείαν για να υπολογιστεί η πιθανότητα $P_m(r)$ για έναν δοσμένο δείκτη r και έναν αριθμό πηγών m.

Μαζί με την πιθανότητα που όλοι οι πόροι χρησιμοποιούνται, μπορεί επίσης να υπολογιστεί ο αριθμός των πελατών που θα πρέπει να παραπεμφθούν σε άλλο σύστημα. Αφού η ζήτηση συνεχίζει με έναν ρυθμό 1/α ανεξάρτητα από την κατάσταση πληρότητας του συστήματος, υπολογίζεται ο ρυθμός ροής ως εξής:

$$\begin{aligned} &\text{Ρυμός ροής} \\ &= \text{Ρυθμός ζήτησης} \\ &* \text{Πιθανότητα να μην είναι απασχολημένοι όλοι οι εξυπηπρετητές} \\ &= \frac{1}{\alpha} * (1 - P_m) \end{aligned} \tag{3.4}$$

Η παραπάνω περίπτωση δίνει άλλο ένα παράδειγμα για το πώς η μεταβλητότητα πρέπει να προσαρμοστεί στην διαδικασία με την τοποθέτηση πλεονάζουσας χωρητικότητας. Για παράδειγμα, ένα επίπεδο αξιοποίησης της τάξεως του 22% σε ένα περιβάλλον με υψηλό σταθερό κόστος είναι ο εφιάλτης κάθε διαχειριστή. Ακόμη, από τη μεριά ενός ατόμου που είναι υπεύθυνο για την δημιουργία μιας διαδικασίας που ανταποκρίνεται γρήγορα, οι απόλυτοι αριθμοί αξιοποίησης θα πρέπει να αντιμετωπίζονται με προσοχή. Ένα κύριο πλεονέκτημα του τύπου που αναφέρθηκε παραπάνω είναι ότι μπορεί γρήγορα να γίνει η αξιολόγηση του πώς οι αλλαγές στη διαδικασία επηρεάζουν την πληρότητα του συστήματος. Για παράδειγμα, μπορεί να υπολογιστεί η πιθανότητα της πληρότητας η οποία θα ήταν αποτέλεσμα μια αυξανόμενης αξιοποίησης. Ένας τέτοιος υπολογισμός θα ήταν σημαντικός, τόσο για να προβλεφθεί η συχνότητες της πληρότητας, όσο και ο ρυθμό ροής.

Η εικόνα 3.1 δείχνει τη σχέση μεταξύ του επιπέδου της έμμεσης αξιοποίησης(implied utilization) και της πιθανότητας η διαδικασία να μην μπορεί να δεχτεί άλλες αφίξεις. Όπως φαίνεται, όπως και με τα προβλήματα αναμονής υπάρχουν σημαντικές οικονομίες κλίμακας και στο σύστημα απώλειας: καθώς μία αξιοποίηση της τάξης του 50% θα οδηγούσε σε πιθανότητα πληρότητας της τάξεως του 30% με έναν εξυπηρετητή (m=1), οδηγεί σε μόνο 13% πιθανότητα πληρότητας με τρεις εξυπηρετητές και λιγότερο από 2% με 10 εξυπηρετητές.



Εικόνα 3.1 Έμμεση αξιοποίηση vs. Πιθανότητα όλοι οι εξυπηρετητές να χρησιμοποιούνται (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

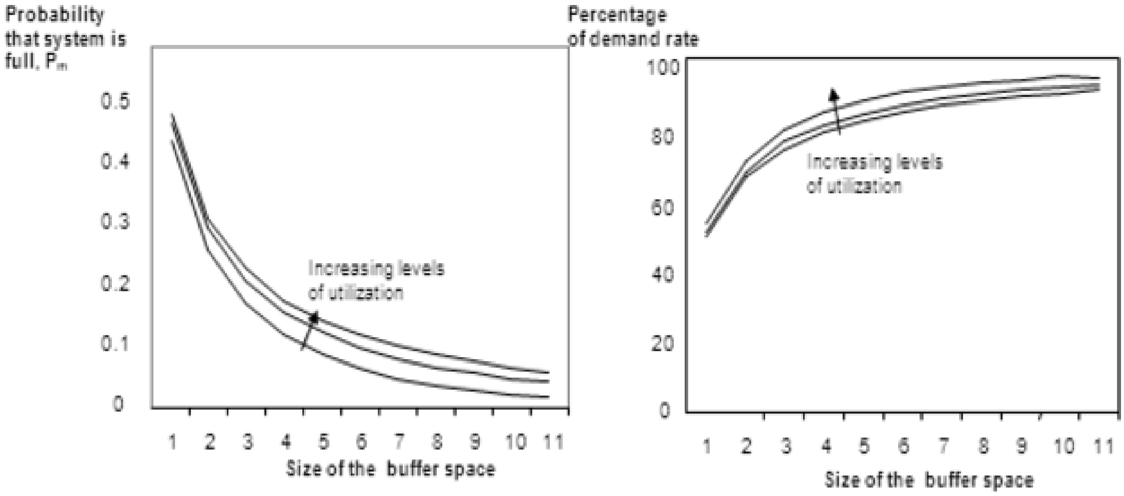
3.4 Η ανυπομονησία των μονάδων ροής (πελατών) και η απόδοση της απώλειας

Στο προηγούμενο κεφάλαιο αναλύθηκε μια διαδικασία στην οποία οι μονάδες ροής περίμεναν υπομονετικά στη σειρά τους μέχρι να εξυπηρετηθούν. Αντίθετα, στο κεφάλαιο αυτό παρατηρείται ότι οι μονάδες ροής δεν περίμεναν και όταν όλοι οι εξυπηρετητές ήταν απασχολημένοι, μετατρέπονταν αμέσως σε χαμένες μονάδες ροής. Αυτές οι δύο περιπτώσεις, ένα πρόβλημα αναμονής στη μία και ένα πρόβλημα απώλειας στην άλλη, είναι σημαντικές, αλλά επίσης είναι ακραίες περιπτώσεις όσον αφορά το αντίκτυπο της μεταβλητότητας στην απόδοση της διαδικασίας. Πολλές ενδιαφέρουσες εφαρμογές που μπορεί να συναντήσετε βρίσκονται κάπου στη μέση αυτών των δύο άκρων. Χωρίς να γίνει περεταίρω ανάλυση, είναι σημαντικό τουλάχιστον να γίνει μία αναφορά σε αυτές τις ενδιάμεσες περιπτώσεις από εννοιολογικής πλευράς.

Μία απ' αυτές τις περιπτώσεις είναι όταν υπάρχει αποθηκευτικός χώρος με περιορισμένο χώρο αναμονής, μπορεί δηλαδή να εξυπηρετηθεί ένας συγκεκριμένος αριθμός μονάδων. Το όριο του μεγέθους του αποθηκευτικός χώρος μπορεί να παρουσιάζει μια απ' τις ακόλουθες περιπτώσεις:

- Σε ένα τηλεφωνικό κέντρο, υπάρχει ένας συγκεκριμένος αριθμός κλήσεων που μπορούν να είναι ταυτόχρονα σε αναμονή. Οι πελάτες που τηλεφωνούν την ώρα που όλες οι γραμμές είναι απασχολημένες λαμβάνουν ένα σήμα κατειλημμένου.
- Ως αποτέλεσμα από το νόμο του Little, ένα όριο στο μέγεθος της ουράς μπορεί απλά να αντιπροσωπεύει ένα μέγιστο χρόνο που είναι διατεθειμένοι οι πελάτες να περιμένουν, αφού ουσιαστικά ο αριθμός των πελατών μας δείχνει και τον αναμενόμενο χρόνο αναμονής.

Για την περίπτωση τους ενός εξυπηρετητή, η εικόνα 3.2 δείχνει την σχέση μεταξύ του αριθμού των διαθέσιμων αποθηκευτικών χώρων και την πιθανότητα όλοι οι αποθηκευτικοί χώροι να είναι γεμάτοι. Αυτή είναι η πιθανότητα η διαδικασία να μην μπορεί να δεχθεί άλλους πελάτες. Όπως είναι φανερό, η πιθανότητα μειώνεται ραγδαία όσο προστίθεται περισσότερος αποθηκευτικός χώρος. Ας σημειωθεί ότι το γράφημα μετατοπίζεται όσο αυξάνονται τα επίπεδα αξιοποίησης.



Εικόνα 3.2 Το αντίκτυπο του αποθηκευτικού χώρου στην πιθανότητα Pm (Πηγή: Cachon- Terwiesch ''Matching Supply with Demand'')

Η απόδοση του συστήματος υπολογίζεται ως εξής:
(1-Probability that all buffers are full)* Demand rate

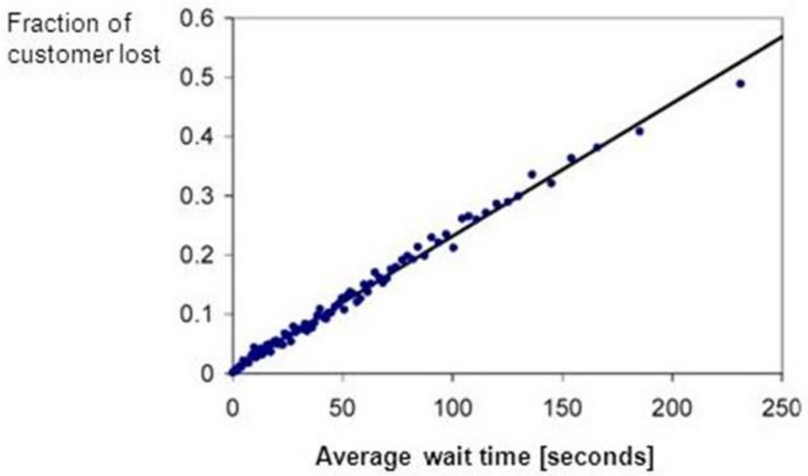
Η δεξιά μεριά της εικόνας 3.2 δείχνει το αντίκτυπο του μεγέθους αποθηκευτικού χώρου στην απόδοση. Ακόμα και για έναν εξυπηρετητή και αξιοποίηση της τάξεως του 90% ,χρειάζονται περισσότεροι από 10 αποθηκευτικοί χώροι για να γίνει η προσέγγιση της αποκατάστασης της απόδοσης που ήταν αναμενόμενη με την έλλειψη της μεταβλητότητας.

Οι δύο ενδιάμεσες περιπτώσεις που αναφέρθηκαν πιο πάνω μοιάζουν μεταξύ τους αλλά διαφέρουν στο ότι στην δεύτερη περίπτωση ο πελάτης εισέρχεται πάντα στο σύστημα αλλά κάποια στιγμή την εγκαταλείπει γιατί κουράστηκε να περιμένει. . Ο τεχνικός όρος γι' αυτήν την περίπτωση είναι «οι πελάτες *εγκαταλείπουν* την ουρά(customers *abandon* the queue)» ή οι πελάτες δειλιάζουν (the customers balk). Αυτή η περίπτωση είναι συνηθισμένη σε τηλεφωνικά κέντρα με μεγάλους χρόνους αναμονής (π. χ τηλεφωνικά κέντρα τραπεζών).

Φαίνονται τώρα στην παρακάτω εικόνα τα δεδομένα ενός τηλεφωνικού κέντρου (τα οποία συλλέχθηκαν από τους Gans, Koole και Mandelbaum (2003)) με μεγάλους χρόνους αναμονής. Ο οριζόντιος άξονας δείχνει πόση ώρα πρέπει να περιμένουν οι πελάτες μέχρι να μιλήσουν με κάποιον πράκτορα. Ο κάθετος άξονας παριστάνει το ποσοστό των πελατών που κλείνουν το τηλέφωνο χωρίς να έχουν εξυπηρετηθεί. Παρατηρείται ότι όσο αυξάνεται ο χρόνος αναμονής τόσο αυξάνονται και οι πελάτες που εγκαταλείπουν το σύστημα.

Υπάρχουν τρεις τρόποι για να βελτιωθούν οι δύο ενδιάμεσες περιπτώσεις, οι περιορισμένοι αποθηκευτικοί χώροι και οι πελάτες που εγκαταλείπουν:

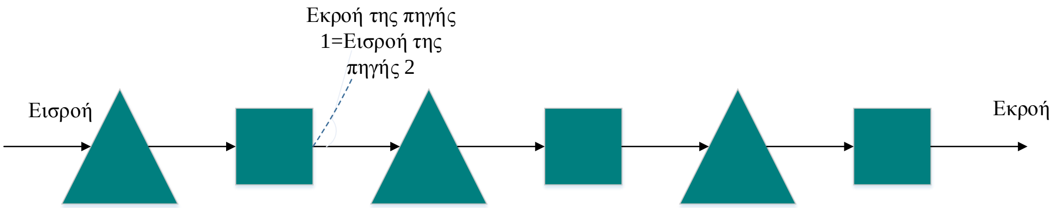
- Μείωση των χρόνων αναμονής. Παρόμοια με την προηγούμενη ανάλυση, οτιδήποτε μπορεί να γίνει για να μειωθούν οι χρόνοι αναμονής (έξυπνη επιλογή χωρητικότητας, μείωση της μεταβλητότητας κλπ) βοηθάει στη μείωση της απόδοσης της απώλειας που προκύπτει από την ανυπομονησία των πελατών.
- Αύξηση του αριθμού των μονάδων ροής που μπορούν να εισέρχονται στον αποθηκευτικό χώρο. Αυτό μπορεί να γίνει είτε αυξάνοντας την προθυμία των πελατών να περιμένουν είτε αυξάνοντας τον αποθηκευτικό χώρο.
- Αποφυγή της εγκατάλειψης των πελατών που είναι ήδη σε αναμονή. Είναι προτιμότερο να φεύγουν οι πελάτες κατευθείαν παρά να περιμένουν και να εγκαταλείπουν μετά το σύστημα. Ο καλύτερος τρόπος για να μην εγκαταλείπει ένας πελάτης το σύστημα επειδή βαρέθηκε είναι να ανακοινώνεται ο εκτιμώμενος χρόνος αναμονής (όπως κάνουν στην Forthnet). Σε αυτή την περίπτωση θα ξέρει ο πελάτης από την αρχή αν θέλει να περιμένει ή αν θέλει να κάνει κάποια άλλη δουλειά την ώρα που περιμένει.



Εικόνα 3.3 Το αντίκτυπο του χρόνου αναμονής στην απώλεια πελατών (Πηγή: Cachon-Terwiesch "Matching Supply with Demand")

3.5 Πηγές με σειριακή μεταβλητότητα

Αφού αναλύθηκε το αντίκτυπο της μεταβλητότητας στην περίπτωση όπου η διαδικασία έχει μία πηγή, τώρα αναφέρεται το αντίκτυπό της σε μία διαδικασία με πολλές πηγές όπως φαίνεται στην εικόνα 3.4. Τέτοιες διαδικασίες είναι πολύ συνηθισμένες, τόσο στο περιβάλλον της παραγωγής όσο και των υπηρεσιών. Στο κομμάτι της παραγωγής ένα παράδειγμα αποτελεί η συναρμολόγηση ενός ποδηλάτου (αποτελείται από πολλά στάδια). Σαν παράδειγμα μιας διαδικασίας παροχής υπηρεσιών που αποτελείται από πολλές πηγές στη σειρά, σκεφτείτε τη διαδικασία της μετανάστευσης στα περισσότερα αεροδρόμια. Όταν φτάνουν ταξιδιώτες στις Ηνωμένες Πολιτείες πρώτα πρέπει να περάσουν από την μεταναστευτική αρχή και μετά να παραταχθούν στο τελωνείο.



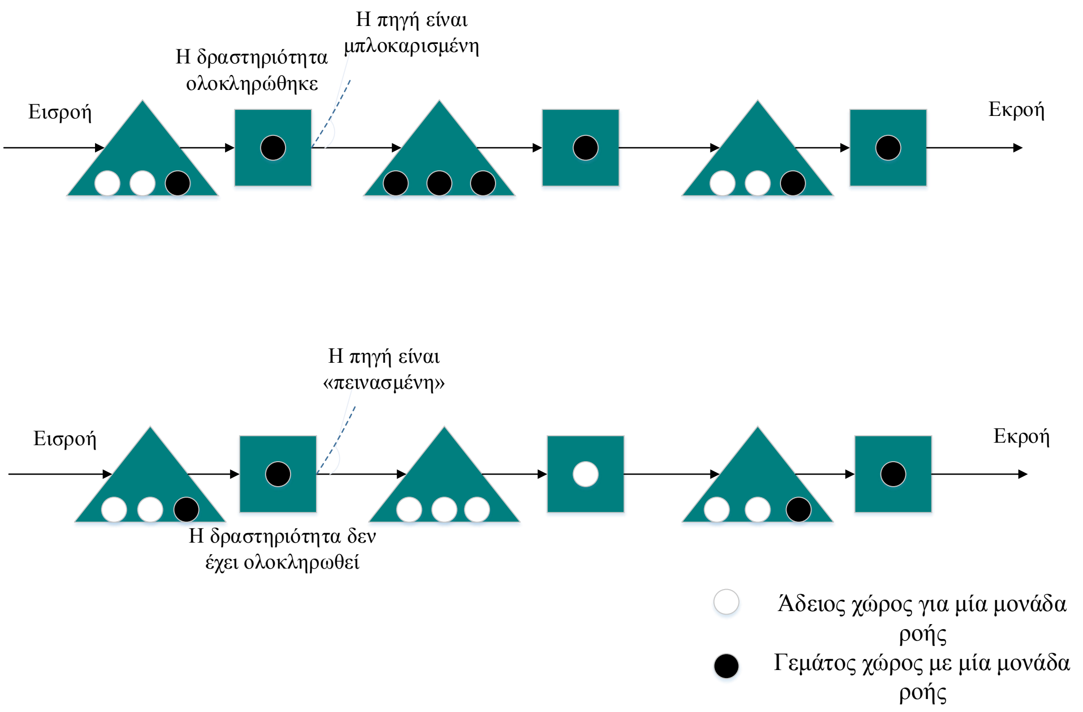
Εικόνα 3.4 Σειριακό σύστημα αναμονής με τρεις πηγές (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Ένας παράγοντας που περιπλέκει τα πράγματα στην ανάλυσή τέτοιων διαδικασιών είναι ότι οι επακόλουθες πηγές δεν δρουν ανεξάρτητα η μία από την άλλη: Η διαδικασία αναχώρησης της πρώτης πηγής είναι η διαδικασία άφιξης της δεύτερης πηγής κοκ. Έτσι, η μεταβλητότητα της διαδικασίας άφιξης της δεύτερης πηγής εξαρτάται από τη μεταβλητότητα της διαδικασίας άφιξης της πρώτης πηγής και από την μεταβλητότητα της διαδικασίας παροχής υπηρεσιών της πρώτης πηγής.

Ανεξάρτητα από την ικανότητά μας να χειριστούμε τέτοιες διαδικασίες, που αναφέρονται σαν διαδοχικές ουρές, εισάγονται κάποια βασικά στοιχεία για το πώς συμπεριφέρονται τέτοιες διαδικασίες.

Ο ρόλος των αποθηκευτικών χώρων

Οι αποθηκευτικοί χώροι έχουν τη δυνατότητα να βελτιώνουν το ρυθμό ροής(flow rate) μέσω της διαδικασίας. Ενώ, στην περίπτωση του ενός εξυπηρετητή, οι αποθηκευτικοί χώροι αυξάνουν τον ρυθμό ροής καθώς μειώνουν την πιθανότητα οι μονάδες που εισέρχονται να μην μπορούν να μπουν στο σύστημα, το αντίκτυπο των αποθηκευτικών χώρων στις διαδοχικές ουρές(tandem queues) είναι κάπως πιο περίπλοκο. Κοιτώντας την εικόνα παρακάτω μπορούμε να διακρίνουμε δύο πράγματα που οδηγούν στη μείωση του ρυθμού ροής σε μία διαδοχική ουρά. Μία πηγή μπορεί να είναι είτε μπλοκαρισμένη είτε «πεινασμένη» (starved). Μπλοκαρισμένη αν δεν είναι σε θέση να απελευθερώσει τη μονάδα ροής που έχει ολοκληρωθεί, καθώς δεν υπάρχει διαθέσιμος αποθηκευτικός χώρος στην επόμενη πηγή και «πεινασμένη» αν είναι σε αδράνεια και ο αποθηκευτικός χώρος που την τροφοδοτεί είναι άδειος.

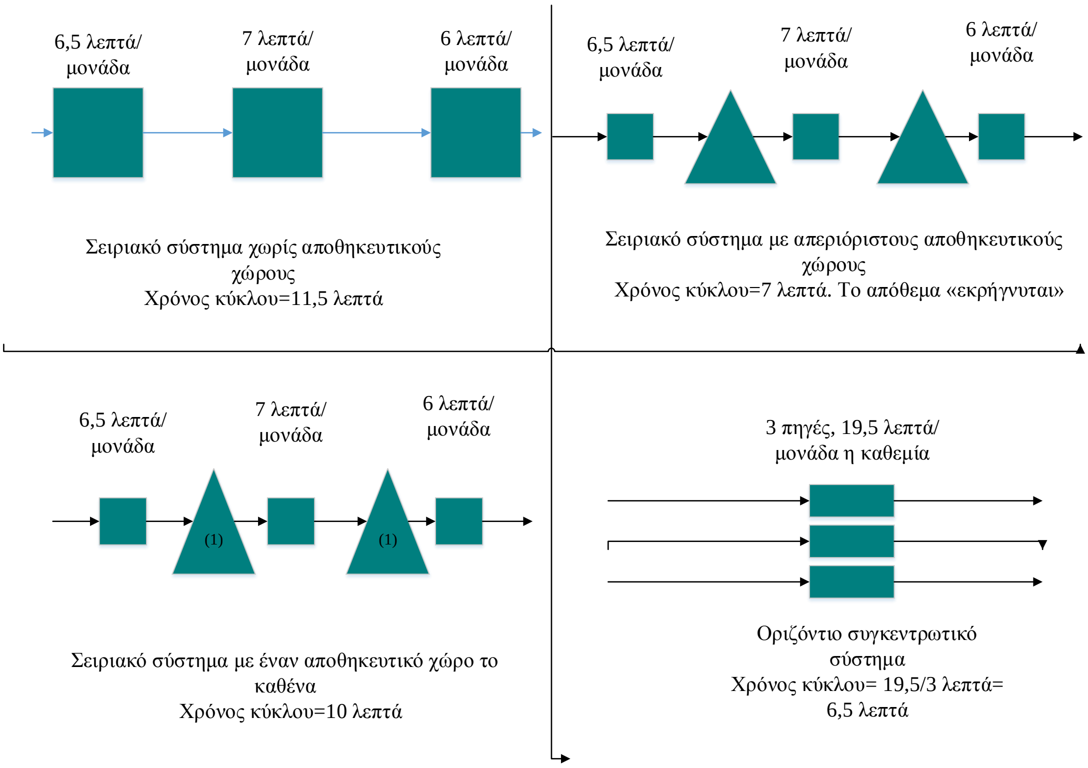


Εικόνα 3.5 Οι έννοιες του μπλοκαρίσματος και της "πείνας" (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Λίγο παραπάνω αναφέρθηκε η πληρότητα ενός συστήματος, για την οποία η σημαντικότερη αιτία είναι το μπλοκάρισμα. Για παράδειγμα, στα επείγοντα ενός νοσοκομείου, ένας ασθενής εκτός από την ώρα που χρειάζεται να εξυπηρετηθεί, θα πρέπει να περιμένει και κάποια ακόμα ώρα μέχρι να βρεθεί διαθέσιμο κρεβάτι. Και έτσι αφού σε αυτήν την περίπτωση η πτέρυγα του νοσοκομείου δεν μπορεί να χρησιμοποιηθεί από κάποιον καινούριο ασθενή, η γεμάτη εντατική μπλοκάρει το νοσοκομείο.

Το μπλοκάρισμα και η "πείνα" μπορούν να αποφευχθούν με την προσθήκη χώρων αποθήκευσης. Οι χώροι αποθήκευσης θα πρέπει να περιλαμβάνουν έναν επαρκή αριθμό μονάδων ροής ώστε να αποφευχθεί η "πείνα" μιας μεταγενέστερης πηγής. Επίσης, θα πρέπει να υπάρχει και αρκετός χώρος ώστε να μην μπλοκάρει η προγενέστερη πηγή. Μερικά νοσοκομεία πειραματίστηκαν εισάγοντας κάποια δωμάτια που προορίζονται για εκείνους του ασθενείς που είναι έτοιμοι να πάνε σπίτι: Ακόμα και ένας επιπλέον χώρος στο τέλος της διαδικασίας (υγιής ασθενής), θα μειώσει την πιθανότητα να μεταφερθεί ένας καινούριος ασθενής λόγω πληρότητας των επειγόντων.
Μαζί με την πιθανότητα να μην μπορούν να γίνουν δεκτές καινούριες μονάδες ροής, ένα εξίσου σημαντικό μέτρο απόδοσης για την διαδικασία μας είναι ο ρυθμός ροής. Η εικόνα 3.5 χρησιμοποιεί μια προσομοίωση για να συγκρίνει τέσσερα σχεδιαγράμματα μιας διαδικασίας τριών πηγών με μεταβλητότητα. Αυτό ανταποκρίνεται σε μία γραμμή του ρυθμού εργασίας, με έναν εργάτη σε κάθε πηγή. Οι χρόνοι επεξεργασίας είναι εκθετικά κατανεμημένοι με μέση τιμή 6.5 λεπτά/μονάδα., 7 λεπτά/μονάδα και 6 λεπτά/μονάδα αντίστοιχα.

Βάσει των μέσων όρων, αναμένεται να παράγεται μία μονάδα κάθε επτά λεπτά. Ωστόσο, με την απουσία χώρου αποθήκευσης, η διαδικασία παράγει μία μονάδα ανά 11.5 λεπτά (επάνω αριστερά). Η διαδικασία δεν αντιλαμβάνεται την ανεπαρκή χωρητικότητα, καθώς η είσοδος είναι συχνά μπλοκαρισμένη (ο σταθμός 2 έχει ολοκληρώσει μία μονάδα ροής όμως δεν μπορεί να την προωθήσει στο σταθμό 3) ή "πεινασμένη" (ο σταθμός 2 θέλει να ξεκινήσει την παραγωγή μιας καινούριας μονάδας αλλά δεν λαμβάνει αποτελέσματα από τον προγενέστερο σταθμό).



Εικόνα 3.6 Ρυθμός ροής σε σύγκριση με τέσσερις διαμορφώσεις ενός συστήματος αναμονής (Πηγή: Cachon- Terwiesch "Matching Supply with Demand" Τροποποιημένος)

Αν εισαχθούν αποθηκευτικοί χώροι σε αυτή τη διαδικασία, ο ρυθμός ροής βελτιώνεται. Απλά και μόνο επιτρέποντας μία μονάδα στον αποθηκευτικό χώρο πριν και μετά το σημείο συμφόρησης αυξάνεται το αποτέλεσμα σε μία μονάδα κάθε 10 λεπτά (κάτω αριστερά). Όμως, παρατηρείται ότι ο αποθηκευτικός χώρος μεταξύ του πρώτου και του δεύτερου βήματος θα μεγαλώσει πολύ γρήγορα.

Τέλος, το κάτω δεξί μέρος της εικόνας 3.6 δείχνει έναν εναλλακτικό τρόπο να γίνει επαναφορά του ρυθμού ροής, διαφορετικά από τη μέθοδο "buffer or suffer"(να βρεθεί δηλαδή αποθηκευτικός χώρος ή αλλιώς θα υποφέρεις). Συνδυάζοντας τις τρεις δραστηριότητες σε μία, εξαλείφεται το μπλοκάρισμα και η "πείνα". Η μέθοδος λέγεται *οριζόντια συγκέντρωση,* καθώς δείχνει την έννοια της συγκέντρωση πανομοιότυπων δραστηριοτήτων και των προηγούμενών τους ξεχωριστών αφίξεων.

Δεδομένου του κόστους των αποθεμάτων και των αρνητικών επιπτώσεων στην ποιότητα, χρειάζεται προσοχή για το που και πόσο απόθεμα (αποθηκευτικό χώρο) επιτρέπεται στην διαδικασία. Καθώς το σημείο συμφόρησης είναι ο περιορισμός για το ρυθμό ροής μέσα στην διαδικασία, καλό είναι αυτό το σημείο να μην είναι ούτε μπλοκαρισμένο ούτε “πεινασμένο”. Έτσι, οι αποθηκευτικοί χώροι είναι πολύ χρήσιμο να τοποθετηθούν πριν και μετά το σημείο συμφόρησης.

Κεφάλαιο 4: Εφαρμογές της θεωρίας ουρών αναμονής σε: (1) Τηλεφωνικό κέντρο τράπεζας (2) Νοσοκομείο

Αφού διατυπώθηκε και αναλύθηκε η θεωρία των ουρών αναμονής, στο τελευταίο κεφάλαιο της εργασίας παρουσιάζονται δύο εφαρμογές όλων των παραπάνω. Στην πρώτη εφαρμογή αναλύεται ένα τηλεφωνικό κέντρο μίας τράπεζας και παρατηρείται τι γίνεται στις περιπτώσεις με έναν και με δύο τηλεφωνητές. Η δεύτερη εφαρμογή αναφέρεται σε ένα νοσοκομείο και συγκεκριμένα στο Κ.Α.Τ και υπολογίζεται η μέση απώλεια ασθενών τις ώρες των εφημεριών ανάλογα με τον αριθμό των γιατρών που εφημερεύουν.

4.1 Τηλεφωνικό κέντρο τράπεζας

Τα περισσότερα προβλήματα που λύνονται με τη βοήθεια των ουρών αναμονής έχουν ως σκοπό την καλύτερη εξυπηρέτηση των πελατών. Είναι πολύ σημαντικό για τον πελάτη, όταν εισέρχεται σε μία ουρά να περιμένει όσο λιγότερο γίνεται και να έχει την καλύτερη και γρηγορότερη εξυπηρέτηση. Όμως ένα αποδοτικό και γρήγορο σύστημα εξυπηρέτησης είναι κάτι που θέλουν και οι διαχειριστές γιατί αυτό που επιζητούν είναι το μικρότερο κόστος αλλά και την ικανοποίηση του πελάτη.
Με τη βοήθεια του excel θα παρουσιάσουμε ένα σύστημα εξυπηρέτησης σε ένα τηλεφωνικό κέντρο μίας τράπεζας. Θα δούμε δύο συστήματα, το πρώτο με έναν τηλεφωνητή και το δεύτερο με δύο τηλεφωνητές. Τα δύο αυτά συστήματα θα έχουν τα ίδια δεδομένα, εκτός από τον αριθμό των τηλεφωνητών και τους μέσους χρόνους άφιξης και εξυπηρέτησης των πελατών.
Το τηλεφωνικό κέντρο αυτής της τράπεζας έχει και έναν βασικό περιορισμό, ότι λειτουργεί καθημερινά από τις 9:00 το πρωί μέχρι τις 17:00 όλο το χρόνο. Άρα η πρώτη άφιξη των πελατών γίνεται στις 9:00, ενώ το τελευταίο τηλεφώνημα θα είναι το πολύ στις 17:00, αφού αν κάποιος καλέσει μετά από αυτή την ώρα η κλήση του δεν θα απαντηθεί, άρα δεν θα καταγραφεί και για το παράδειγμα παρακάτω. Επίσης υποθέτουμε ότι η ουρά είναι της μορφής FIFO(Fist Come First Served), δηλαδή ο καθένας εξυπηρετείται ανάλογα με τη σειρά που καλεί, και έτσι οι πελάτες δεν φεύγουν λόγω μεγάλης αναμονής. Τέλος, το σύστημα που χρησιμοποιούμε, ακόμα και στην δεύτερη περίπτωση με τους δύο τηλεφωνητές, έχει μόνο μία ουρά από την οποία τροφοδοτείται.
Θα ξεκινήσουμε λοιπόν την υλοποίηση του προβλήματος στο Excel.Το excel είναι ένα πρόγραμμα που αποτελείται από λογιστικά φύλλα και μπορεί να κάνει διάφορους μαθηματικούς υπολογισμούς, απλούς αλλά και πολύπλοκους, να επεξεργάζεται και να δημιουργεί βάσεις δεδομένων και πολλά άλλα. Τα λογιστικά αυτά φύλλα αποτελούνται από γραμμές και στήλες οι οποίες δημιουργούν κελιά μέσα στα οποία μπορούμε να καταχωρήσουμε διάφορα δεδομένα από αριθμούς και κείμενα μέχρι πολύπλοκες μαθηματικές συναρτήσεις. Η υλοποίηση στο excel βασίζεται στη λογική του

εύκολου χειρισμού του μοντέλου από τρίτους που απλά θέλουν να συλλέξουν αποτελέσματα χωρίς να πρέπει να ασχοληθούν με τις συναρτήσεις που κρύβονται πίσω από αυτά τα αποτελέσματα.
Ξεκινάμε αρχικά με το φύλλο Αποτελέσματα όπου για να υπολογίσουμε όλα τα υπόλοιπα θα πρέπει πρώτα να θέσουμε τη μέση τιμή των ενδιάμεσων χρόνων άφιξης όπου θα πούμε ότι είναι 4 λεπτά, όπως και τη μέση τιμή των χρόνων εξυπηρέτησης που έστω ότι είναι 3 λεπτά. Βάζοντας αυτές τις τιμές φαίνεται και ότι οι μέση τιμή των ενδιάμεσων χρόνων αφίξεων είναι μεγαλύτερη από τη μέση τιμή των χρόνων επεξεργασίας. Επίσης για να γίνουν σωστά οι υπολογισμοί θα πρέπει να μορφοποιήσουμε κατάλληλα τα κελιά κάτι που στην περίπτωσή μας σημαίνει ότι θα διαλέξουμε την προσαρμοσμένη μορφοποίηση λλ:δδ (λεπτά και δευτερόλεπτα).

	A	B
1		
2	Είσοδος	
3	Μέση τιμή ενδιάμεσων χρόνων άφιξης	Μέσης τιμή χρόνων εξυπηρέτησης
4	04:00	03:00
5	mm:ss	mm:ss

Εικόνα 4.1 Είσοδος συστήματος

Συνεχίζουμε με το φύλλο Δεδομένα στο οποίο παρουσιάζονται τα δεδομένα για την παραγωγή των αποτελεσμάτων που χρειαζόμαστε. Αρχικά έχουμε τις τέσσερις πρώτες στήλες οι οποίες περιέχουν:
Στήλη Α: Αύξων αριθμός του πελάτη που εισέρχεται στο σύστημα
Στήλη Β: Οι ενδιάμεσοι χρόνοι μεταξύ των αφίξεων οι οποίοι είναι τυχαίοι αριθμοί που προκύπτουν από τη μέση τιμή των ενδιάμεσων χρόνων άφιξης που βάλαμε στο φύλλο Αποτελέσματα. Η συνάρτηση η οποία θα έχουμε είναι η =-Αποτελέσματα!$A$4* LN(RAND()) όπου LN(RAND()) ο φυσικός αλγόριθμος της συνάρτησης RAND που μας δίνει τιμές 0 και 1.Αυτός ο τύπος προκύπτει αφού σε μία διαδικασία Poisson με ρυθμό λ για να προσομοιώσουμε n χρόνους πραγματοποίησης γεγονότων παράγουμε n ανεξάρτητες εκθετικές τυχαίες μεταβλητές Exp(λ) και έχουμε $x_i = -\frac{1}{\lambda} * lnU_i$ όπου U είναι ένας τυχαίος αριθμός και x_i ο ενδιάμεσος χρόνος μεταξύ του i-1 και του i-στου γεγονότος στη διαδικασία Poisson. Είναι σημαντικό να βάλουμε το σήμα «$» στο κελί Α4 για να υπολογίζει πάντα το γινόμενο από το συγκεκριμένο κελί όσο κάτω και αν προχωρήσουμε στη στήλη Β.

B4 =-Αποτελέσματα!$A$4*LN(RAND())

Πελάτης	Χρόνος μεταξύ των αφίξεων	Χρόνος εξυπηρέτησης	Χρόνος άφιξης
			9:00
1	12:50	13:30	9:12
2	00:35	01:20	9:13
3	03:31	00:17	9:16
4	01:05	02:06	9:18
5	02:59	02:17	9:20
6	02:15	00:31	9:23
7	00:24	03:20	9:23
8	00:20	01:02	9:23
9	02:31	03:30	9:26
10	01:28	00:04	9:27
11	00:16	00:25	9:28
12	00:43	04:40	9:28
13	03:42	04:11	9:32

Εικόνα 4.2 Χρόνος μεταξύ των αφίξεων

Στήλη C: Ο χρόνος εξυπηρέτησης όπου χρησιμοποιούμε τον τύπο που είδαμε στη στήλη B αλλά αντί για το κελί A4 τώρα έχουμε το B4, δηλαδή τη μέση τιμή των χρόνων εξυπηρέτησης.

C4 =-Αποτελέσματα!$B$4*LN(RAND())

	A	B	C	D
1–2	Πελάτης	Χρόνος μεταξύ των αφίξεων	Χρόνος εξυπηρέτησης	Χρόνος άφιξης
3				9:00
4	1	12:50	13:30	9:12
5	2	00:35	01:20	9:13
6	3	03:31	00:17	9:16
7	4	01:05	02:06	9:18
8	5	02:59	02:17	9:20
9	6	02:15	00:31	9:23
10	7	00:24	03:20	9:23
11	8	00:20	01:02	9:23
12	9	02:31	03:30	9:26
13	10	01:28	00:04	9:27
14	11	00:16	00:25	9:28
15	12	00:43	04:40	9:28
16	13	03:42	04:11	9:32

Εικόνα 4.3 Χρόνος εξυπηρέτησης

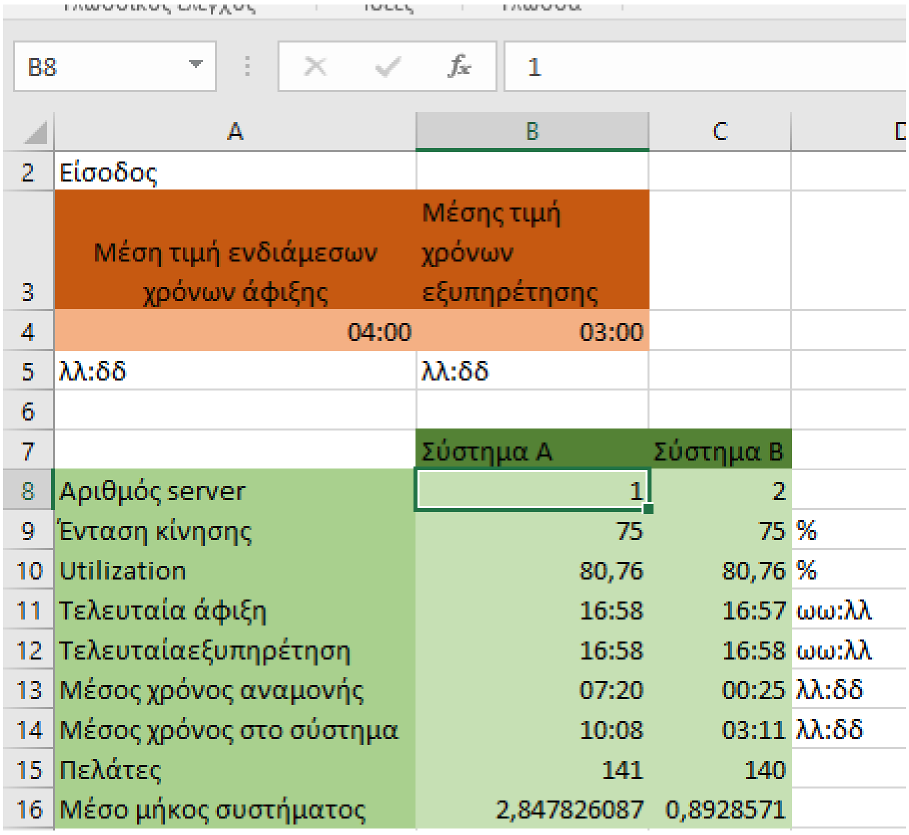
Στήλη D: Ο χρόνος άφιξης του κάθε πελάτη. Εδώ επειδή έχουμε ώρα, αλλάζουμε τη μορφοποίηση του κελιού σε προσαρμοσμένη μορφοποίηση ωω:λλ (ώρες και λεπτά). Θέτουμε μία αρχική τιμή όπου εδώ είναι 9 η ώρα το πρωί εφόσον εκείνη την ώρα ξεκινάει η λειτουργία του τηλεφωνικού κέντρου. Στα επόμενα τώρα κελιά θα υπολογίζουμε τον προηγούμενο χρόνο άφιξης συν το χρόνο μεταξύ των αφίξεων όπως για παράδειγμα στο κελί D4 έχουμε D3+B4 (βλ. Εικόνα 4.1)

D4 =D3+B4

	A	B	C	D
1–2	Πελάτης	Χρόνος μεταξύ των αφίξεων	Χρόνος εξυπηρέτησης	Χρόνος άφιξης
3				9:00
4	1	12:50	13:30	9:12
5	2	00:35	01:20	9:13
6	3	03:31	00:17	9:16
7	4	01:05	02:06	9:18
8	5	02:59	02:17	9:20
9	6	02:15	00:31	9:23
10	7	00:24	03:20	9:23
11	8	00:20	01:02	9:23
12	9	02:31	03:30	9:26
13	10	01:28	00:04	9:27
14	11	00:16	00:25	9:28
15	12	00:43	04:40	9:28
16	13	03:42	04:11	9:32

Εικόνα 4.4 Χρόνος άφιξης

Στις επόμενες στήλες E με N έχουμε δεδομένα από τα δύο συστήματα, τα οποία σχετίζονται άμεσα με το φύλλο Αποτελέσματα.



B8 | f_x | 1

	A	B	C	D
2	Είσοδος			
3	Μέση τιμή ενδιάμεσων χρόνων άφιξης	Μέσης τιμή χρόνων εξυπηρέτησης		
4	04:00	03:00		
5	λλ:δδ	λλ:δδ		
6				
7		Σύστημα Α	Σύστημα Β	
8	Αριθμός server	1	2	
9	Ένταση κίνησης	75	75	%
10	Utilization	80,76	80,76	%
11	Τελευταία άφιξη	16:58	16:57	ωω:λλ
12	Τελευταίαεξυπηρέτηση	16:58	16:58	ωω:λλ
13	Μέσος χρόνος αναμονής	07:20	00:25	λλ:δδ
14	Μέσος χρόνος στο σύστημα	10:08	03:11	λλ:δδ
15	Πελάτες	141	140	
16	Μέσο μήκος συστήματος	2,847826087	0,8928571	

Εικόνα 4.5 Το φύλλο Αποτελέσματα

Στις στήλες E μέχρι και I θα δούμε τα δεδομένα του συστήματος Α. Στη στήλη E βλέπουμε την Αρχή της εξυπηρέτησης η οποία υπολογίζεται από τον τύπο :

=MAX(D4;IF(A4>=Αποτελέσματα!$B$8;LARGE($F$3:F3;Αποτελέσματα!$B$8);0))

Όπου η συνάρτηση MAX μας δίνει τον μεγαλύτερο από τους αριθμούς που της δίνουμε. Σ' αυτήν την περίπτωση της δώσαμε την τιμή του κελιού D4 και τη συνάρτηση IF. Η συνάρτηση IF μας δίνει μία τιμή αν η συνθήκη που θα δώσουμε είναι αληθής και μία άλλη αν η συνθήκη είναι ψευδής. Στην περίπτωσή μας θέλουμε να δούμε αν ο αύξων αριθμός του πελάτη είναι μεγαλύτερος από τον αριθμό των εξυπηρετητών που έχουμε στο σύστημα, κάτι που το βλέπουμε στο κελί B8 του φύλλου Αποτελέσματα (βλ. Εικόνα4.5). Αν η συνθήκη είναι αληθής μας δίνει τη συνάρτηση LARGE. Η συνάρτηση LARGE μας δίνει τη ν-στή μεγαλύτερη τιμή μεταξύ ενός εύρους τιμών, εδώ τις τιμές της στήλης F, από την αρχή της εξυπηρέτησης μέχρι και τον τρέχοντα χρόνο λήξης της εξυπηρέτησης. Αν η συνθήκη IF είναι ψευδής τότε παίρνουμε την τιμή 0.

E4 =MAX(D4;IF(A4>=Αποτελέσματα!$B$8;LARGE($F$3:F3;Αποτελέσματα!$B$8);0))

	E	F	G	H	I
1	Εξυπηρέτηση συστήματος Α				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	00:00	13:30	0
5	9:26	9:27	12:56	14:16	1
6	9:27	9:27	10:45	11:02	2
7	9:27	9:30	09:57	12:04	3
8	9:30	9:32	09:05	11:22	4
9	9:32	9:32	09:07	09:39	5
10	9:32	9:36	09:14	12:34	6
11	9:36	9:37	12:14	13:16	7
12	9:37	9:40	10:45	14:15	7
13	9:40	9:40	12:47	12:51	7
14	9:40	9:41	12:35	13:00	7
15	9:41	9:45	12:17	16:57	8
16	9:45	9:50	13:15	17:26	7

Εικόνα 4.6 Αρχή εξυπηρέτησης

Η στήλη F είναι το άθροισμα της αρχής της εξυπηρέτησης με τον αντίστοιχο χρόνο εξυπηρέτησης (στήλες Ε και C). Αυτή η στήλη είναι απαραίτητη για τη δημιουργία της προηγούμενης και επίσης πρέπει να της δώσουμε μία τιμή αρχικοποίησης.

F4 =E4+C4

	E	F	G	H	I
1	Εξυπηρέτηση συστήματος Α				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	00:00	13:30	0
5	9:26	9:27	12:56	14:16	1
6	9:27	9:27	10:45	11:02	2
7	9:27	9:30	09:57	12:04	3
8	9:30	9:32	09:05	11:22	4
9	9:32	9:32	09:07	09:39	5
10	9:32	9:36	09:14	12:34	6
11	9:36	9:37	12:14	13:16	7
12	9:37	9:40	10:45	14:15	7
13	9:40	9:40	12:47	12:51	7
14	9:40	9:41	12:35	13:00	7
15	9:41	9:45	12:17	16:57	8
16	9:45	9:50	13:15	17:26	7

Εικόνα 4.7 Τέλος εξυπηρέτησης

Η στήλη G είναι πολύ σημαντική και παρουσιάζει το χρόνο που περίμενε ο πελάτης από την ώρα που κάλεσε μέχρι και να αρχίσει η εξυπηρέτησή του. Αυτός ο χρόνος αναμονής είναι η διαφορά της εκάστοτε τιμής της στήλης Ε με την αντίστοιχη τιμή της στήλης D.

G4 =E4-D4

	E	F	G	H	I
1	Εξυπηρέτηση συστήματος Α				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	00:00	13:30	0
5	9:26	9:27	12:56	14:16	1
6	9:27	9:27	10:45	11:02	2
7	9:27	9:30	09:57	12:04	3
8	9:30	9:32	09:05	11:22	4
9	9:32	9:32	09:07	09:39	5
10	9:32	9:36	09:14	12:34	6
11	9:36	9:37	12:14	13:16	7
12	9:37	9:40	10:45	14:15	7
13	9:40	9:40	12:47	12:51	7
14	9:40	9:41	12:35	13:00	7
15	9:41	9:45	12:17	16:57	8
16	9:45	9:50	13:15	17:26	7

Εικόνα 4.8 Χρόνος αναμονής

Η στήλη H μας δείχνει το χρόνο που περνάει ο πελάτης στο σύστημα από την ώρα που κάλεσε μέχρι και το τέλος της εξυπηρέτησής του, όπου εδώ είναι η διαφορά της στήλης F(τέλος εξυπηρέτησης) με τη στήλη D(χρόνος άφιξης του πελάτη).

H4 =F4-D4

	E	F	G	H	I
1	Εξυπηρέτηση συστήματος Α				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	00:00	13:30	0
5	9:26	9:27	12:56	14:16	1
6	9:27	9:27	10:45	11:02	2
7	9:27	9:30	09:57	12:04	3
8	9:30	9:32	09:05	11:22	4
9	9:32	9:32	09:07	09:39	5
10	9:32	9:36	09:14	12:34	6
11	9:36	9:37	12:14	13:16	7
12	9:37	9:40	10:45	14:15	7
13	9:40	9:40	12:47	12:51	7
14	9:40	9:41	12:35	13:00	7
15	9:41	9:45	12:17	16:57	8
16	9:45	9:50	13:15	17:26	7

Εικόνα 4.9 Χρόνος στο σύστημα

Η τελευταία στήλη για το σύστημα Α μας δείχνει πόσα άτομα περιμένουν κάθε φορά στην ουρά. Χρησιμοποιούμε τη συνάρτηση COUNTIF η οποία μετράει τα κελιά της στήλης F από της αρχικοποίηση μέχρι και το πέρας της εξυπηρέτησης του προηγούμενου πελάτη, τα οποία έχουν μεγαλύτερη τιμή από την άφιξη του εκάστοτε πελάτη.

I4 =COUNTIF($F$3:F3;">"&D4)

	E	F	G	H	I
1	Εξυπηρέτηση συστήματος Α				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	00:00	13:30	0
5	9:26	9:27	12:56	14:16	1
6	9:27	9:27	10:45	11:02	2
7	9:27	9:30	09:57	12:04	3
8	9:30	9:32	09:05	11:22	4
9	9:32	9:32	09:07	09:39	5
10	9:32	9:36	09:14	12:34	6
11	9:36	9:37	12:14	13:16	7
12	9:37	9:40	10:45	14:15	7
13	9:40	9:40	12:47	12:51	7
14	9:40	9:41	12:35	13:00	7
15	9:41	9:45	12:17	16:57	8
16	9:45	9:50	13:15	17:26	7

Εικόνα 4.10 Μήκος συστήματος

Όπως είδαμε για τις στήλες E μέχρι και I, αντίστοιχα υπολογίζονται και οι στήλες J μέχρι και N με τα αντίστοιχα δεδομένα για το σύστημα Β (βλ. Εικόνα 4.2).Είναι σημαντικό να αναφέρουμε πως και τα δύο συστήματα παίρνουν τους ίδιους χρόνους άφιξης, εξυπηρέτησης και τους ενδιάμεσους χρόνους.

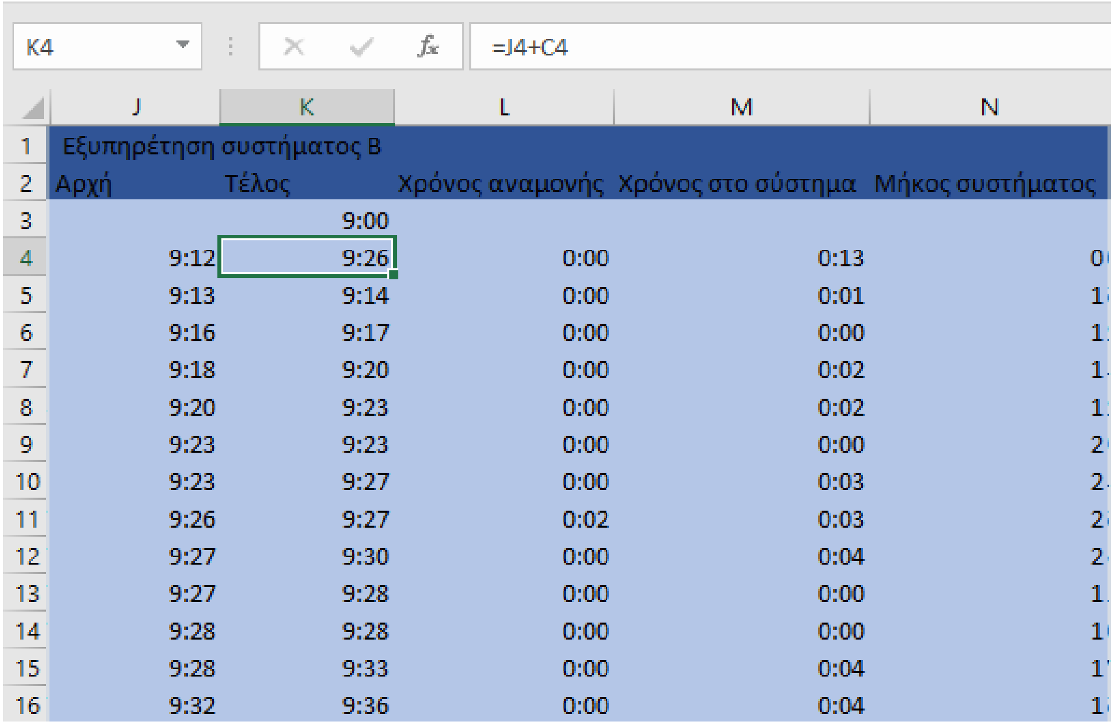
J4 =MAX(D4;IF(A4>=Αποτελέσματα!$C$8;LARGE($K$3:K3;Αποτελέσματα!$C$8);0))

	J	K	L	M	N
1	Εξυπηρέτηση συστήματος Β				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	0:00	0:13	0
5	9:13	9:14	0:00	0:01	1
6	9:16	9:17	0:00	0:00	1
7	9:18	9:20	0:00	0:02	1
8	9:20	9:23	0:00	0:02	1
9	9:23	9:23	0:00	0:00	2
10	9:23	9:27	0:00	0:03	2
11	9:26	9:27	0:02	0:03	2
12	9:27	9:30	0:00	0:04	2
13	9:27	9:28	0:00	0:00	1
14	9:28	9:28	0:00	0:00	1
15	9:28	9:33	0:00	0:04	1
16	9:32	9:36	0:00	0:04	1

Εικόνα 4.11 Σύστημα Β (Αρχή εξυπηρέτησης)

Εδώ όπως και στο σύστημα Α έχουμε τη συνάρτηση =MAX(D4;IF(A4>=Αποτελέσματα!$C$8;LARGE($K$3:K3;Αποτελέσματα!$C$8);0)) όμως εδώ έχουμε το κελί C8 αφού είμαστε στο σύστημα με τους δύο τηλεφωνητές και επίσης αντλούμε τα δεδομένα μας από τη στήλη K.

Συνεχίζοντας είναι η στήλη K είναι το άθροισμα της αρχής της εξυπηρέτησης με τον αντίστοιχο χρόνο εξυπηρέτησης (στήλες J και C).Όπως και στο σύστημα Α, έτσι και εδώ θέτουμε μία τιμή αρχικοποίησης (εδώ 9:00).



	J	K	L	M	N
1	Εξυπηρέτηση συστήματος Β				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	0:00	0:13	0
5	9:13	9:14	0:00	0:01	1
6	9:16	9:17	0:00	0:00	1
7	9:18	9:20	0:00	0:02	1
8	9:20	9:23	0:00	0:02	1
9	9:23	9:23	0:00	0:00	2
10	9:23	9:27	0:00	0:03	2
11	9:26	9:27	0:02	0:03	2
12	9:27	9:30	0:00	0:04	2
13	9:27	9:28	0:00	0:00	1
14	9:28	9:28	0:00	0:00	1
15	9:28	9:33	0:00	0:04	1
16	9:32	9:36	0:00	0:04	1

Εικόνα 4.12 Σύστημα Β (Τέλος εξυπηρέτησης)

Η στήλη L είναι ο χρόνος που περίμενε ο πελάτης από την ώρα που κάλεσε μέχρι και να αρχίσει η εξυπηρέτησή του. Αυτός ο χρόνος αναμονής είναι η διαφορά της εκάστοτε τιμής της στήλης J με την αντίστοιχη τιμή της στήλης D.

L4 =J4-D4

	J	K	L	M	N
1	Εξυπηρέτηση συστήματος Β				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	0:00	0:13	0
5	9:13	9:14	0:00	0:01	1
6	9:16	9:17	0:00	0:00	1
7	9:18	9:20	0:00	0:02	1
8	9:20	9:23	0:00	0:02	1
9	9:23	9:23	0:00	0:00	2
10	9:23	9:27	0:00	0:03	2
11	9:26	9:27	0:02	0:03	2
12	9:27	9:30	0:00	0:04	2
13	9:27	9:28	0:00	0:00	1
14	9:28	9:28	0:00	0:00	1
15	9:28	9:33	0:00	0:04	1
16	9:32	9:36	0:00	0:04	1

Εικόνα 4.13 Σύστημα Β (Χρόνος αναμονής)

Η στήλη M, όπως και πριν, είναι ο χρόνος που περνάει ο πελάτης στο σύστημα και είναι η διαφορά της στήλης K(τέλος εξυπηρέτησης) με τη στήλη D(χρόνος άφιξης του πελάτη).

M4 =K4-D4

	J	K	L	M	N
1	Εξυπηρέτηση συστήματος Β				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	0:00	0:13	0
5	9:13	9:14	0:00	0:01	1
6	9:16	9:17	0:00	0:00	1
7	9:18	9:20	0:00	0:02	1
8	9:20	9:23	0:00	0:02	1
9	9:23	9:23	0:00	0:00	2
10	9:23	9:27	0:00	0:03	2
11	9:26	9:27	0:02	0:03	2
12	9:27	9:30	0:00	0:04	2
13	9:27	9:28	0:00	0:00	1
14	9:28	9:28	0:00	0:00	1
15	9:28	9:33	0:00	0:04	1
16	9:32	9:36	0:00	0:04	1

Εικόνα 4.14 Σύστημα Β (Χρόνος στο σύστημα)

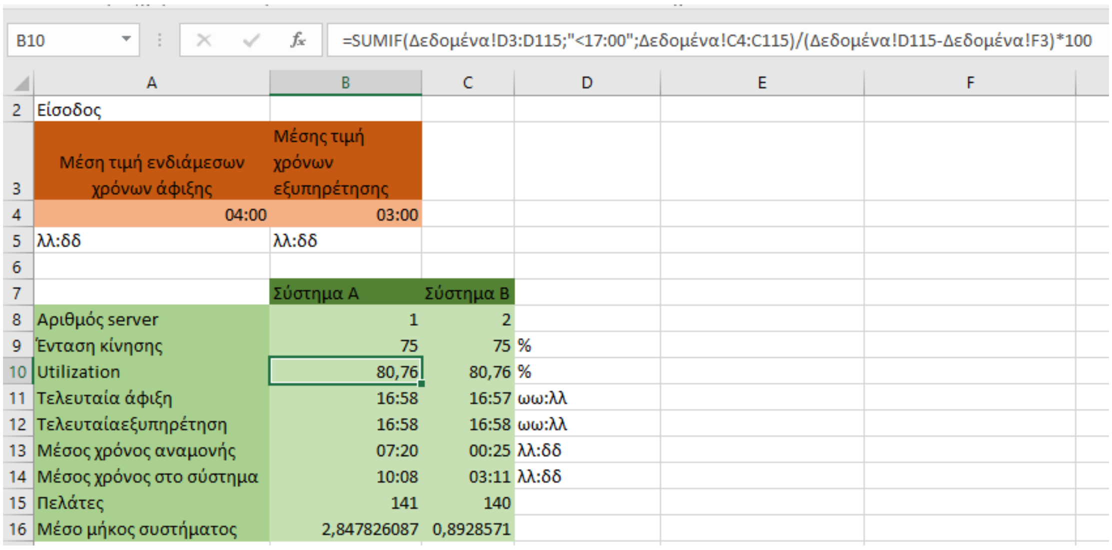
Και τέλος έχουμε το μήκος του συστήματος Β όπου και πάλι χρησιμοποιούμε τη συνάρτηση COUNTIF η οποία μετράει τα κελιά της στήλης K από της αρχικοποίηση μέχρι και το πέρας της εξυπηρέτησης του προηγούμενου πελάτη, τα οποία έχουν μεγαλύτερη τιμή από την άφιξη του εκάστοτε πελάτη.

N4 =COUNTIF($K$3:K3;">"&D4)

	J	K	L	M	N
1	Εξυπηρέτηση συστήματος Β				
2	Αρχή	Τέλος	Χρόνος αναμονής	Χρόνος στο σύστημα	Μήκος συστήματος
3		9:00			
4	9:12	9:26	0:00	0:13	0
5	9:13	9:14	0:00	0:01	1
6	9:16	9:17	0:00	0:00	1
7	9:18	9:20	0:00	0:02	1
8	9:20	9:23	0:00	0:02	1
9	9:23	9:23	0:00	0:00	2
10	9:23	9:27	0:00	0:03	2
11	9:26	9:27	0:02	0:03	2
12	9:27	9:30	0:00	0:04	2
13	9:27	9:28	0:00	0:00	1
14	9:28	9:28	0:00	0:00	1
15	9:28	9:33	0:00	0:04	1
16	9:32	9:36	0:00	0:04	1

Εικόνα 4.15 Σύστημα Β (Μήκος συστήματος)

Αφού είδαμε όλα τα δεδομένα και τους υπολογισμούς μας, θα ξαναγυρίσουμε στο φύλλο Αποτελέσματα, αφού αυτό είναι όπως είπαμε το φύλλο που θα δει ο αναλυτής για να συλλέξει τα δεδομένα και τις πληροφορίες που χρειάζεται.
Με τα πρώτα πεδία του φύλλου ασχοληθήκαμε πιο πάνω και είναι οι μέσες τιμές των ενδιάμεσων χρόνων άφιξης και των χρόνων εξυπηρέτησης.
Προχωρώντας παρακάτω βλέπουμε την 8^{η} γραμμή στην οποία έχουμε τον αριθμό των εξυπηρετητών που έχουμε στο σύστημα. Αυτό τον αριθμό τον βάζουμε εμείς. Παρακάτω έχουμε την ένταση κίνησης του συστήματος. Η ένταση κίνησης είναι το πηλίκο της μέσης τιμής του χρόνου εξυπηρέτησης προς τη μέση τιμή των ενδιάμεσων χρόνων άφιξης επί 100. Η γραμμή νούμερο 9 μας δίνει την αξιοποίηση του συστήματος και εκφράζεται σαν ποσοστό όπως και η ένταση κίνησης παραπάνω.
Η αξιοποίηση είναι το άθροισμα των χρόνων εξυπηρέτησης προς τη διαφορά της λήξης του πρώτου χρόνου εξυπηρέτησης από τον τελευταίο, επί 100. Για να είναι όμως απολύτως σωστό αυτό το όρισμα θα πρέπει να βάλουμε και τον περιορισμό του ωραρίου αφού ξέρουμε πως η υπηρεσία λειτουργεί μέχρι τις 17:00. Έτσι θα βάλουμε μία SUMIF ώστε τελικά στο άθροισμα να υπολογίσει μόνο τους πελάτες εκείνους που έφτασαν μέχρι τις 17:00.

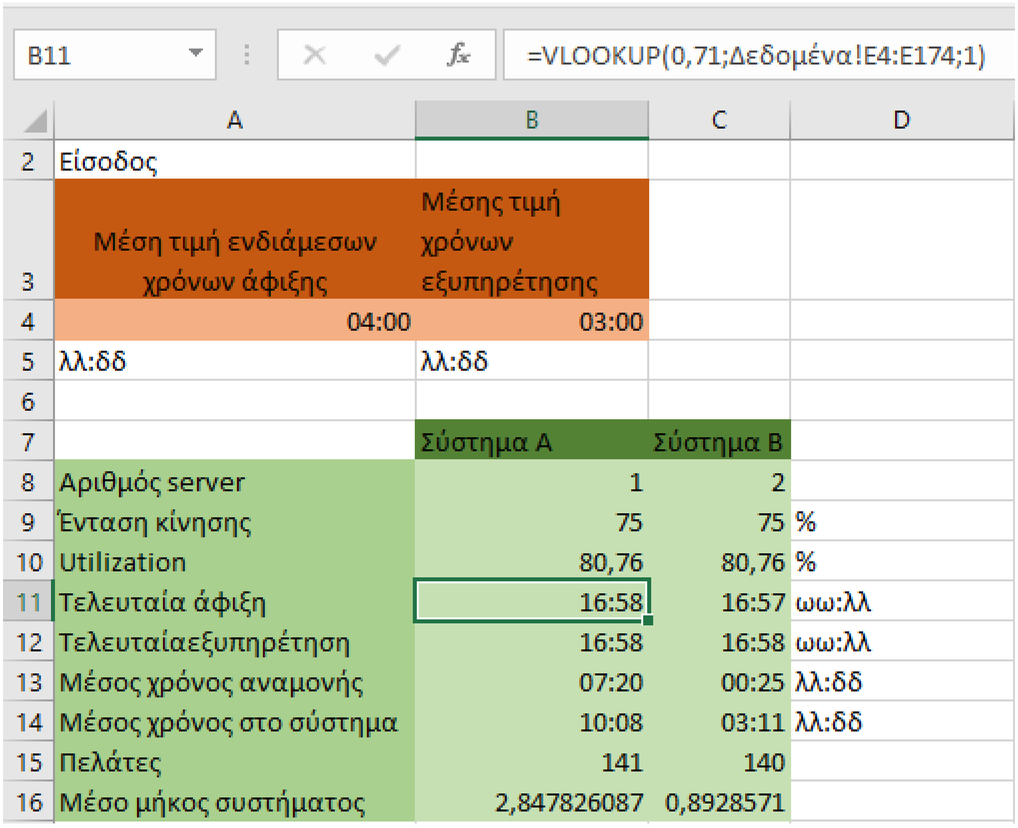


B10 =SUMIF(Δεδομένα!D3:D115;"<17:00";Δεδομένα!C4:C115)/(Δεδομένα!D115-Δεδομένα!F3)*100

	A	B	C	D	E	F
2	Είσοδος					
3	Μέση τιμή ενδιάμεσων χρόνων άφιξης	Μέσης τιμή χρόνων εξυπηρέτησης				
4	04:00	03:00				
5	λλ:δδ	λλ:δδ				
6						
7		Σύστημα Α	Σύστημα Β			
8	Αριθμός server	1	2			
9	Ένταση κίνησης	75	75	%		
10	Utilization	80,76	80,76	%		
11	Τελευταία άφιξη	16:58	16:57	ωω:λλ		
12	Τελευταίαεξυπηρέτηση	16:58	16:58	ωω:λλ		
13	Μέσος χρόνος αναμονής	07:20	00:25	λλ:δδ		
14	Μέσος χρόνος στο σύστημα	10:08	03:11	λλ:δδ		
15	Πελάτες	141	140			
16	Μέσο μήκος συστήματος	2,847826087	0,8928571			

Εικόνα 4.16 Σύγκριση δύο συστημάτων (Utilization)

Στην παρακάτω γραμμή βλέπουμε την ώρα της τελευταίας άφιξης για κάθε σύστημα. Εδώ χρησιμοποιήσαμε τη VLOOKUP η οποία ψάχνει σε μία στήλη ενός πίνακα και μας δίνει την τιμή που του ζητήσαμε. Αν δεν βρει αυτή την τιμή μας δίνει την αμέσως μικρότερη. Έτσι μπορούμε εύκολα να βρούμε την ώρα της τελευταίας άφιξης σε καθένα από τα δύο συστήματα βάζοντας όμως πάντα τον περιορισμό ότι όλες οι τιμές πρέπει να είναι πριν από τις 17:00 (όπου λόγω της επεξεργασίας των κελιών που έχουμε βάλει μέχρι τώρα, σε μορφοποίηση απλού αριθμού θα είναι 0,71).



B11 =VLOOKUP(0,71;Δεδομένα!E4:E174;1)

	A	B	C	D
2	Είσοδος			
3	Μέση τιμή ενδιάμεσων χρόνων άφιξης	Μέσης τιμή χρόνων εξυπηρέτησης		
4	04:00	03:00		
5	λλ:δδ	λλ:δδ		
6				
7		Σύστημα Α	Σύστημα Β	
8	Αριθμός server	1	2	
9	Ένταση κίνησης	75	75	%
10	Utilization	80,76	80,76	%
11	Τελευταία άφιξη	16:58	16:57	ωω:λλ
12	Τελευταίαεξυπηρέτηση	16:58	16:58	ωω:λλ
13	Μέσος χρόνος αναμονής	07:20	00:25	λλ:δδ
14	Μέσος χρόνος στο σύστημα	10:08	03:11	λλ:δδ
15	Πελάτες	141	140	
16	Μέσο μήκος συστήματος	2,847826087	0,8928571	

Εικόνα 4.17 Σύγκριση δύο συστημάτων (Τελευταία άφιξη)

Την ίδια συνάρτηση χρησιμοποιούμε και παρακάτω για να βρούμε την ώρα της τελευταίας εξυπηρέτησης. Εδώ όμως χρησιμοποιούμε τη στήλη που μας δείχνει το τέλος της εξυπηρέτησης και πάλι με τον περιορισμό ότι όλες οι τιμές πρέπει να είναι πριν τις 17:00.
Στη συνέχεια έχουμε το μέσο χρόνο αναμονής που είναι αν όχι το πιο σημαντικό αποτέλεσμα, ένα από τα σημαντικότερα. Υπολογίζεται χρησιμοποιώντας τη συνάρτηση AVERAGEIF η οποία μας δίνει το μέσο όρος της στήλης που θα της δώσουμε με τον περιορισμό όμως η λήξη της λειτουργίας της ουράς να είναι μέχρι τις 17:00. Όπως και με το μέσο χρόνο αναμονής, πολύ σημαντικό μέγεθος είναι και ο μέσος χρόνος στο σύστημα ο οποίος υπολογίζεται και αυτός με την παραπάνω συνάρτηση αλλά μετράει το μέσο όρο της στήλης που μας δίνει το Χρόνο στο σύστημα (στήλες H και M).

B13 fx =AVERAGEIF(Δεδομένα!E4:E174;"<=17:00";Δεδομένα!G4:G174)

	A	B	C	D	E
2	Είσοδος				
3	Μέση τιμή ενδιάμεσων χρόνων άφιξης	Μέσης τιμή χρόνων εξυπηρέτησης			
4	04:00	03:00			
5	λλ:δδ	λλ:δδ			
6					
7		Σύστημα Α	Σύστημα Β		
8	Αριθμός server	1	2		
9	Ένταση κίνησης	75	75	%	
10	Utilization	80,76	80,76	%	
11	Τελευταία άφιξη	16:58	16:57	ωω:λλ	
12	Τελευταίαεξυπηρέτηση	16:58	16:58	ωω:λλ	
13	Μέσος χρόνος αναμονής	07:20	00:25	λλ:δδ	
14	Μέσος χρόνος στο σύστημα	10:08	03:11	λλ:δδ	
15	Πελάτες	141	140		
16	Μέσο μήκος συστήματος	2,847826087	0,8928571		

Εικόνα 4.18 Σύγκριση δύο συστημάτων (Μέσος χρόνος αναμονής)

Έπειτα έχουμε τους πελάτες. Εδώ έχουμε την συνάρτηση COUNTIF η οποία μετράει του πελάτες που μπήκαν στο σύστημα μέχρι τις 17:00 όπου και κλείνει το τηλεφωνικό κέντρο.
Τέλος, έχουμε το μέσο μήκος του συστήματος το οποίο υπολογίζεται και αυτό με τη συνάρτηση AVERAGEIF επιστρέφοντάς μας το αποτέλεσμα σε μορφή αριθμού και όχι ώρας.
Έτσι λοιπόν με τη χρήση του excel υπολογίσαμε τα δύο βασικά μεγέθη που μας απασχόλησαν στο δεύτερο κεφάλαιο, το μέσο χρόνο αναμονής και το μέσο χρόνο στο σύστημα για το παράδειγμα ενός τηλεφωνικού κέντρου μιας τράπεζας και για τις περιπτώσεις με έναν εξυπηρετητή αλλά και με περισσότερους.

4.2 Νοσοκομείο

Στο προηγούμενο παράδειγμα είδαμε την εφαρμογή του μοντέλου που αναλύσαμε στο κεφάλαιο 2. Εδώ θα δούμε ένα παράδειγμα του μοντέλου που περιγράψαμε στο κεφάλαιο 3.

Ο τομέας της υγείας είναι γεμάτος καθυστερήσεις. Πόσες φορές έχει τύχει στον καθένα να περιμένει βδομάδες για να κλείσει ένα ραντεβού με έναν γιατρό και ακόμα και όταν φτάνει η μέρα του ραντεβού να χρειάζεται να περιμένει λίγο ακόμα μέχρι να μπορέσει ο γιατρός να τον εξετάσει. Στα νοσοκομεία κυρίως είναι το πιο συνηθισμένο να βλέπει κανείς δεκάδες ανθρώπους να περιμένουν στα επείγοντα μέχρι να έρθει η σειρά τους να εξετασθούν από τον εφημερεύοντα ιατρό. Όλα αυτά είναι αποτέλεσμα της μεταβλητότητας όπως αναλύθηκε στα παραπάνω κεφάλαια. Παρακάτω θα παρουσιάσουμε ένα παράδειγμα αυτής της αναμονής στα νοσοκομεία αλλά και την πιθανότητα να είναι όλοι οι εφημερεύοντες ιατροί απασχολημένοι, κάτι που θα οδηγήσει στο να χαθούν κάποιοι ασθενείς επειδή δεν έχουν την υπομονή ή και το χρόνο για να περιμένουν.

Τα δεδομένα που θα δούμε είναι από το νοσοκομείο Κ.Α.Τ και είναι παρμένα μερικά από το ίντερνετ και μερικά από κάποιες νοσοκόμες που συνομιλήσαμε γιατί δεν ήταν δυνατή η παροχή τους εγγράφως λόγω απορρήτου του νοσοκομείου.

Τα στοιχεία που συλλέξαμε από το ίντερνετ αναφέρονται στο έτος 2002 και στην Β γενική χειρουργική πτέρυγα του νοσοκομείου που είναι μέρος του Β’ χειρουργικού τομέα. Στον τομέα αυτόν επίσης υπάρχουν και η Α’ γενική χειρουργική, η Α’ ΜΕΘ, Β’ ΜΕΘ, η Νευροχειρουργική, η πλαστική, η μονάδα εγκαυμάτων, η στοματική και γναθοπροσωπική χειρουργική, η αγγειοχειρουργική, η θωρακοχειρουργική, το ουρολογικό και το οδοντιατρικό ιατρείο.

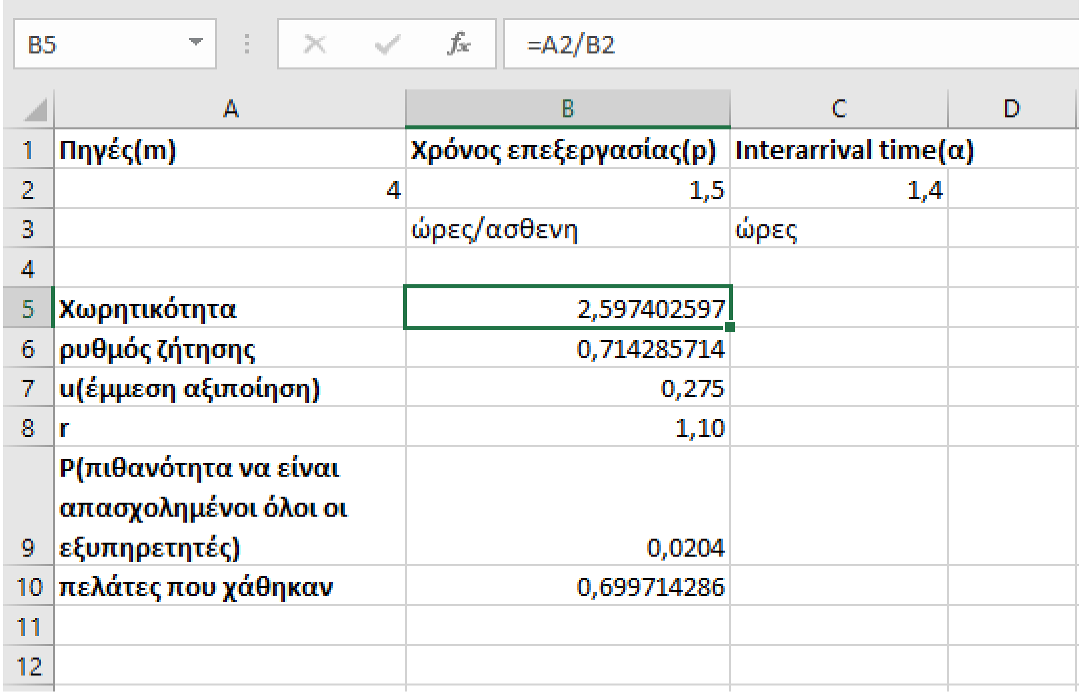
Στα επείγοντα της Β’ γενικής χειρουργικής για το 2002 εξετάστηκαν περίπου 3200 ασθενείς αμιγώς με προβλήματα γενικής χειρουργικής ή με σύνθετα προβλήματα πολλών ειδικοτήτων, συνήθως πολυτραυματίες. Σε κάθε εφημερία παρίστανται 3 ειδικευόμενοι ιατροί και 1-2 επιμελητές. Τέλος, οι νοσοκόμες μας πληροφόρησαν ότι κάθε ασθενής χρειάζεται περίπου 1,5 ώρες από την ώρα που θα εισέλθει μέχρι και την ώρα που θα εξέλθει από το νοσοκομείο αν υπάρχει 1 επιμελητής μαζί με του ειδικευόμενους, ενώ χρειάζονται 1,3 ώρες για τον ασθενή αν παρευρίσκονται 2 επιμελητές εκτός από τους ειδικευόμενους.

Έχοντας λοιπόν τα παραπάνω δεδομένα θα δούμε δύο περιπτώσεις. Στην πρώτη θα εξετάσουμε πόσοι πελάτες θα μπορούσαν να χαθούν αν είχαμε 4 άτομα ανά εφημερία (m=4 πηγές) και στην δεύτερη θα δούμε τι θα γινόταν αν ήταν πάντα 5 άτομα στην εφημερία (m=5 πηγές).

Αρχικά όμως θα πρέπει να υπολογίσουμε κάθε πόση ώρα εισέρχεται ασθενής στα επείγοντα. Αφού είχαμε 3200 ασθενείς όλη τη χρονιά, υπολογίζοντας ότι η εφημερίες είναι κάθε μέρα 12 ώρες περίπου τη μέρα, υπολογίζουμε ότι φτάνει ένας νέος ασθενής κάθε 1,4 ώρες (interarrival time α).

Ας πάρουμε λοιπόν την πρώτη περίπτωση με τους 4 γιατρούς ανά εφημερία. Έχουμε έναν ασθενή κάθε 1,4 ώρες και χρόνο επεξεργασίας 1.5 ώρες.

Αρχικά θα υπολογίσουμε τη χωρητικότητα που έχει το σύστημα υποθέτοντας ότι και οι 4 γιατροί είναι ελεύθεροι. Η χωρητικότητα είναι το πηλίκο των πηγών προς το χρόνο επεξεργασίας. Όπως παρατηρούμε (Εικόνα 4.15) εδώ έχουμε χωρητικότητα 2.7 ασθενείς την ώρα.

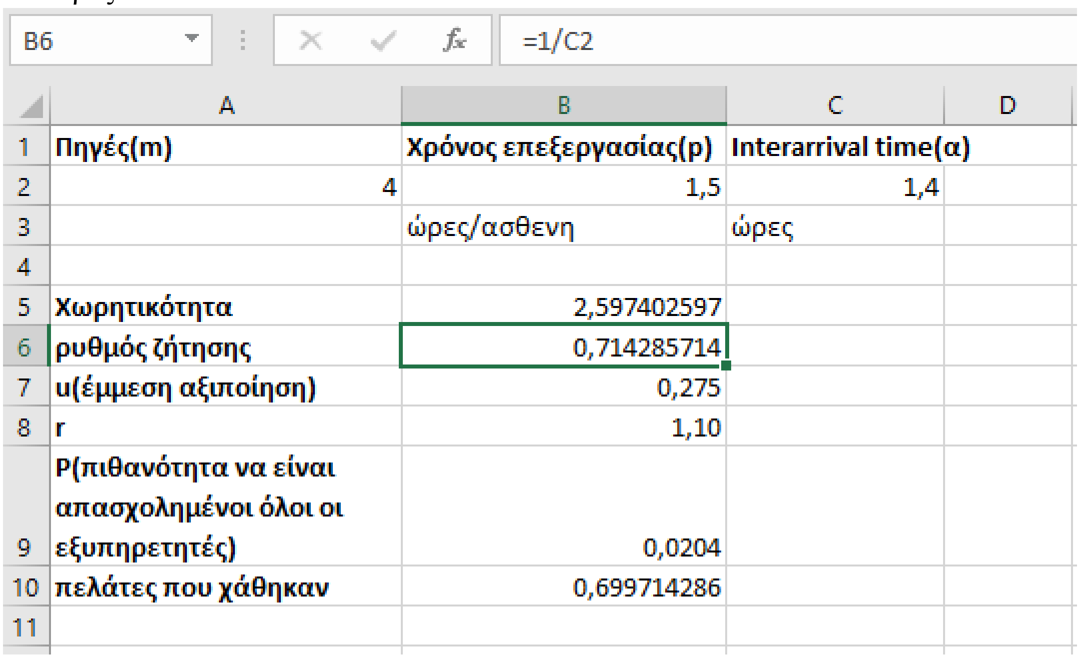


B5 =A2/B2

	A	B	C	D
1	Πηγές(m)	Χρόνος επεξεργασίας(p)	Interarrival time(α)	
2	4	1,5	1,4	
3		ώρες/ασθενη	ώρες	
4				
5	Χωρητικότητα	2,597402597		
6	ρυθμός ζήτησης	0,714285714		
7	u(έμμεση αξιποίηση)	0,275		
8	r	1,10		
9	P(πιθανότητα να είναι απασχολημένοι όλοι οι εξυπηρετητές)	0,0204		
10	πελάτες που χάθηκαν	0,699714286		
11				
12				

Εικόνα 4.19 Χωρητικότητα

Στη συνέχεια υπολογίζουμε το ρυθμό ζήτησης δηλαδή κάθε πόση ώρα φτάνει κάποιος ασθενής, ο οποίος είναι 1 προς το χρόνο μεταξύ των αφίξεων, όπου εδώ είναι 1.4 ώρες.



B6 =1/C2

	A	B	C	D
1	Πηγές(m)	Χρόνος επεξεργασίας(p)	Interarrival time(α)	
2	4	1,5	1,4	
3		ώρες/ασθενη	ώρες	
4				
5	Χωρητικότητα	2,597402597		
6	ρυθμός ζήτησης	0,714285714		
7	u(έμμεση αξιποίηση)	0,275		
8	r	1,10		
9	P(πιθανότητα να είναι απασχολημένοι όλοι οι εξυπηρετητές)	0,0204		
10	πελάτες που χάθηκαν	0,699714286		
11				

Εικόνα 4.20 Ρυθμός ζήτησης

Παρακάτω θα υπολογίσουμε την έμμεση αξιοποίηση η οποία είναι το πηλίκο του ρυθμού ζήτησης προς τη χωρητικότητα.

B7 =B6/B5

	A	B	C	D
1	Πηγές(m)	Χρόνος επεξεργασίας(p)	Interarrival time(α)	
2	4	1,5	1,4	
3		ώρες/ασθενη	ώρες	
4				
5	Χωρητικότητα	2,597402597		
6	ρυθμός ζήτησης	0,714285714		
7	u(έμμεση αξιποίηση)	0,275		
8	r	1,10		
9	P(πιθανότητα να είναι απασχολημένοι όλοι οι εξυπηρετητές)	0,0204		
10	πελάτες που χάθηκαν	0,699714286		
11				
12				

Εικόνα 4.21 Έμμεση αξιοποίηση

Στη συνέχεια θα πρέπει να βρούμε την πιθανότητα (P) να είναι απασχολημένοι όλοι οι εξυπηρετητές. Όμως για να βρούμε αυτή την πιθανότητα θα πρέπει να χρησιμοποιήσουμε τον τύπο απωλειών του Erlang (όπου αναλύθηκε στο κεφάλαιο 3), ο οποίος ορίζει ότι πρέπει να βρούμε πρώτα το r όπου ορίζεται σαν το πηλίκο του χρόνου επεξεργασίας προς το διάστημα μεταξύ των αφίξεων ή ως το γινόμενο της έμμεσης αξιοποίησης με τον αριθμό των πηγών.

B8 =B7*A2

	A	B	C	D
1	Πηγές(m)	Χρόνος επεξεργασίας(p)	Interarrival time(α)	
2	4	1,5	1,4	
3		ώρες/ασθενη	ώρες	
4				
5	Χωρητικότητα	2,597402597		
6	ρυθμός ζήτησης	0,714285714		
7	u(έμμεση αξιποίηση)	0,275		
8	r	1,10		
9	P(πιθανότητα να είναι απασχολημένοι όλοι οι εξυπηρετητές)	0,0204		
10	πελάτες που χάθηκαν	0,699714286		
11				

Εικόνα 4.22 Δείκτης r

Για να βρούμε τώρα την πιθανότητα να είναι όλοι οι εξυπηρετητές απασχολημένοι θα πρέπει να χρησιμοποιήσουμε τον πίνα απώλεια του Erlang. Αυτό που θα κάνουμε είναι πρώτα να βρούμε την αντίστοιχη γραμμή κεφαλίδας (r=1,1) που δείχνει την αναλογία του χρόνου επεξεργασίας προς τα διαστήματα μεταξύ των αφίξεων (βλέπε πίνακα 4.1). Ύστερα, θα βρούμε την αντίστοιχη κεφαλίδα στήλης (m=4) που δείχνει τον αριθμό των πηγών.

Πίνακας 4.1 Πίνακας απώλειας του Erlang (Excel)

Erlang loss table											
m											
		1	2	3	4	5	6	7	8	9	10
	0,1	0,0909	0,0045	0,0002	0	0	0	0	0	0	0
	0,2	0,1667	0,0164	0,0011	0,0001	0	0	0	0	0	0
	0,25	0,2	0,0244	0,002	0,0001	0	0	0	0	0	0
	0,3	0,2308	0,0335	0,0033	0,0003	0	0	0	0	0	0
	0,33	0,25	0,04	0,0044	0,0004	0	0	0	0	0	0
	0,4	0,2857	0,0541	0,0072	0,0007	0,0001	0	0	0	0	0
	0,5	0,3333	0,0769	0,0127	0,0016	0,0002	0	0	0	0	0
	0,6	0,375	0,1011	0,0198	0,003	0,0004	0	0	0	0	0
	0,67	0,4	0,1176	0,0255	0,0042	0,0006	0,0001	0	0	0	0
	0,7	0,4118	0,126	0,0286	0,005	0,0007	0,0001	0	0	0	0
	0,75	0,4286	0,1385	0,0335	0,0062	0,0009	0,0001	0	0	0	0
	0,8	0,4444	0,1509	0,0387	0,0077	0,0012	0,0002	0	0	0	0
r	0,9	0,4737	0,1757	0,0501	0,0111	0,002	0,0003	0	0	0	0
	1	0,5	0,2	0,0625	0,0154	0,0031	0,0005	0,0001	0	0	0
	1,1	0,5238	0,2237	0,0758	0,0204	0,0045	0,0008	0,0001	0	0	0
	1,2	0,5455	0,2466	0,0898	0,0262	0,0063	0,0012	0,0002	0	0	0
	1,25	0,5556	0,2577	0,097	0,0294	0,0073	0,0015	0,0003	0	0	0
	1,3	0,5652	0,2687	0,1043	0,0328	0,0085	0,0018	0,0003	0,0001	0	0
	1,33	0,5714	0,2759	0,1092	0,0351	0,0093	0,0021	0,0004	0,0001	0	0
	1,4	0,5833	0,2899	0,1192	4	0,0111	0,0026	0,0005	0,0001	0	0
	1,5	0,6	0,3103	0,1343	0,048	0,0142	0,0035	0,0008	0,0001	0	0
	1,6	0,6154	0,3299	0,1496	0,0565	0,0177	0,0047	0,0011	0,0002	0	0
	1,67	0,625	0,3425	0,1598	0,0624	0,0204	0,0056	0,0013	0,0003	0,0001	0
	1,7	0,6296	0,3486	0,165	0,0655	0,0218	0,0061	0,0015	0,0003	0,0001	0
	1,75	0,6364	0,3577	0,1726	0,0702	0,024	0,0069	0,0017	0,0004	0,0001	0
	1,8	0,6429	0,3665	0,1803	0,075	0,0263	0,0078	0,002	0,0005	0,0001	0
	1,9	0,6552	0,3836	0,1955	0,085	0,0313	0,0098	0,0027	0,0006	0,0001	0
	2	0,6667	0,4	0,2105	0,0952	0,0367	0,0121	0,0034	0,0009	0,0002	0

Η τομή αυτής της γραμμής με αυτή τη στήλη είναι $P_m(r) = 0.0204$. Στο excel το υπολογίσαμε με τη συνάρτηση INDEX την οποία την βάλαμε να ψάξει σε έναν πίνακα την τομή μιας γραμμής με μία στήλη. Για να βρει όμως την τομή που θέλαμε εμείς χρησιμοποιήσαμε μέσα στην INDEX δύο φορές τη συνάρτηση MATCH για να βρει τα στοιχεία που θέλαμε τη μία στη αντίστοιχη γραμμή και την άλλη στην αντίστοιχη στήλη. Το αποτέλεσμα αυτό που βρήκαμε σε αυτή την περίπτωση είναι ότι το νοσοκομείο θα βρίσκεται σε κατάσταση πληρότητας για το 2%του χρόνου.

B9 =INDEX(J4:S31;MATCH(B8;I4:I31);MATCH(A2;J3:S3))

	A	B	C
1	Πηγές(m)	Χρόνος επεξεργασίας(p)	Interarrival time(α)
2	4	1,5	1,4
3		ώρες/ασθενη	ώρες
4			
5	Χωρητικότητα	2,597402597	
6	ρυθμός ζήτησης	0,714285714	
7	u(έμμεση αξιποίηση)	0,275	
8	r	1,10	
9	P(πιθανότητα να είναι απασχολημένοι όλοι οι εξυπηρετητές)	0,0204	
10	πελάτες που χάθηκαν	0,699714286	

Erlang loss table

		m							
		1	2	3	4	5	6	7	
	0,1	0,0909	0,0045	0,0002	0	0	0	0	0
	0,2	0,1667	0,0164	0,0011	0,0001	0	0	0	0
	0,25	0,2	0,0244	0,002	0,0001	0	0	0	0
	0,3	0,2308	0,0335	0,0033	0,0003	0	0	0	0
	0,33	0,25	0,04	0,0044	0,0004	0	0	0	0
	0,4	0,2857	0,0541	0,0072	0,0007	0,0001	0	0	0
	0,5	0,3333	0,0769	0,0127	0,0016	0,0002	0	0	0
	0,6	0,375	0,1011	0,0198	0,003	0,0004	0	0	0
	0,67	0,4	0,1176	0,0255	0,0042	0,0006	0,0001	0	0
	0,7	0,4118	0,126	0,0286	0,005	0,0007	0,0001	0	0
	0,75	0,4286	0,1385	0,0335	0,0062	0,0009	0,0001	0	0
	0,8	0,4444	0,1509	0,0387	0,0077	0,0012	0,0002	0	0
r	0,9	0,4737	0,1757	0,0501	0,0111	0,002	0,0003	0	0
	1	0,5	0,2	0,0625	0,0154	0,0031	0,0005	0,0001	0
	1,1	0,5238	0,2237	0,0758	0,0204	0,0045	0,0008	0,0001	0
	1,2	0,5455	0,2466	0,0898	0,0262	0,0063	0,0012	0,0002	0
	1,25	0,5556	0,2577	0,097	0,0294	0,0073	0,0015	0,0003	0
	1,3	0,5652	0,2687	0,1043	0,0328	0,0085	0,0018	0,0003	0,00

Εικόνα 4.23 Πιθανότητα όλοι οι εξυπηρετητές να είναι απασχολημένοι

Τέλος, υπολογίζουμε τους πελάτες που χάθηκαν από το σύστημα ως το γινόμενος του ρυθμού ζήτησης επί την πιθανότητα να μην είναι όλοι οι εξυπηρετητές απασχολημένοι (1-$P_m(r)$). Όπως βλέπουμε στην παρακάτω εικόνα στην περίπτωση των 4 πηγών βλέπουμε ότι χάνονται 0,7 περίπου ασθενής την ώρα, δηλαδή περίπου 8 ασθενείς τη μέρα.

B10 =(1/C2)*(1-B9)

	A	B	C
1	Πηγές(m)	Χρόνος επεξεργασίας(p)	Interarrival time(α)
2	4	1,5	1,4
3		ώρες/ασθενη	ώρες
4			
5	Χωρητικότητα	2,597402597	
6	ρυθμός ζήτησης	0,714285714	
7	u(έμμεση αξιποίηση)	0,275	
8	r	1,10	
9	P(πιθανότητα να είναι απασχολημένοι όλοι οι εξυπηρετητές)	0,0204	
10	πελάτες που χάθηκαν	0,699714286	
11			

Εικόνα 4.24 Ασθενείς που χάθηκαν

Αφού παρατηρήσαμε τι θα γινόταν στην περίπτωση των 4 πηγών τώρα θα δούμε και την περίπτωση με τις 5 πηγές (γιατρούς).

B9 =INDEX(Φύλλο1!J4:S31;MATCH(Φύλλο3!B8;Φύλλο1!I4:I31);MATCH(Φύλλο3!A2;Φύλλο1!J3:S3))

	A	B	C	D	E	F	G	H	I
1	**Πηγές(m)**	**Χρόνος επεξεργασίας(p)**	**Interarrival time(α)**						
2	5	1,3	1,4						
3		ώρες/ασθενή	ώρες						
4									
5	**Χωρητικότητα**	3,968253968							
6	**ρυθμός ζήτησης**	0,714285714							
7	**u(έμμεση αξιποίηση)**	0,18							
8	**r**	0,900							
9	**P(πιθανότητα να είναι απασχολημένοι όλοι οι εξυπηρετητές)**	0,0012							
10	**πελάτες που χάθηκαν**	0,713428571							
11									
12									

Εικόνα 4.25 Περίπτωση 5 πηγών

Όπως βλέπουμε στην παραπάνω εικόνα στην περίπτωση αυτή που έχουμε και μικρότερο χρόνο επεξεργασίας, η πιθανότητα να είναι όλοι οι εξυπηρετητές απασχολημένοι είναι αρκετά μικρότερη που σημαίνει ότι το νοσοκομείο θα ήταν σε πληρότητα το 1,2% του χρόνου. Παρόλα αυτά παρατηρούμε ότι οι πελάτες που χάνονται είναι και πάλι περίπου 0,7 την ώρα. Άρα βλέπουμε ότι ο ένας παραπάνω γιατρός δεν θα βοηθούσε τόσο πολύ.

Κεφάλαιο 5: Συμπεράσματα και περαιτέρω έρευνα

Στην παραπάνω εργασία αναλύθηκαν τα συστήματα ουρών αναμονής και έγινε εστιασμός στις υπηρεσίες. Αναφερθήκαν τα δίκτυα ουρών αναμονής, τα χαρακτηριστικά μίας ουράς, αναφέρθηκαν κάποια από τα βασικά μοντέλα ουρών αναμονής και παρουσιάστηκαν κάποιες βασικές κατανομές. Στη συνέχεια παρατηρήθηκε το αντίκτυπο της μεταβλητότητας στις ουρές αναμονής και δόθηκε ιδιαίτερη προσοχή στους χρόνους αναμονής με έναν αλλά και περισσότερους εξυπηρετητές όπως επίσης και στην απώλεια της απόδοσης. Τέλος, φάνηκε στην πράξη, με δύο παραδείγματα στο Excel, πώς η μεταβλητότητα επηρεάζει τους χρόνους αναμονής ανάλογα με τον αριθμό των εξυπηρετητών που διαθέτουμε αλλά το ρόλο που παίζει στην απώλεια των πελατών.
Η θεωρία ουρώ αναμονής ουσιαστικά είναι μία μαθηματική έρευνα των συστημάτων αναμονής. Το πρώτο παράδειγμα, του τηλεφωνικού κέντρου της τράπεζας, είναι μία εφαρμογή αυτής της έρευνας μέσω της οποίας μπορεί να υπολογιστούν παράγοντες όπως το μέσο μήκος της ουράς, τον αναμενόμενο αριθμό πελατών στο σύστημα, το μέσο χρόνο που περνά ένας πελάτης στο σύστημα, και το πιο σημαντικό, το μέσο χρόνο που πρέπει να περιμένει ο πελάτης στο σύστημα μέχρι να εξυπηρετηθεί. Όπως φάνηκε λοιπόν και στο παράδειγμα, ο χρόνος αναμονής επηρεάζεται άμεσα από το χρόνο άφιξης αλλά και από το μέσο χρόνο εξυπηρέτησης. Επίσης, παρατηρήθηκε ότι με την προσθήκη του δεύτερου τηλεφωνητή, σύστημα Β, ο μέσος χρόνος αναμονής αλλά και ο μέσος χρόνος εξυπηρέτησης μειώνεται αισθητά και σχεδόν μηδενίζεται. Αν είχαμε ακόμα ένα σύστημα Γ, με τρεις εξυπηρετητές, ο μέσος χρόνος αναμονής θα πλησίαζε ακόμα περισσότερο το μηδέν. Άρα ακόμα ένας βασικός παράγοντας που επηρεάζει το χρόνο αναμονής είναι και ο αριθμός των εξυπηρετητών που βρίσκονται στο σύστημα.
Η σημαντικότητα των εξυπηρετητών αλλά και των μέσων χρόνων επεξεργασίας και άφιξης φαίνεται και στο δεύτερο παράδειγμα που παρατέθηκε. Σε αυτό το παράδειγμα του νοσοκομείου, στην περίπτωση με τους πέντε γιατρούς σε κάθε εφημερία, παρατηρήθηκε ότι η πιθανότητα να είναι όλοι απασχολημένοι είναι μικρότερη από την περίπτωση όπου οι γιατροί της εφημερίας είναι τέσσερις. Επίσης, φάνηκε ότι στη δεύτερη περίπτωση το νοσοκομείο βρίσκεται σε κατάσταση πληρότητας το 2% του χρόνου, ενώ με τους τέσσερις γιατρούς, όπου και ο χρόνος εξυπηρέτησης είναι μεγαλύτερος, βρίσκεται σε πληρότητα το 1,2% του χρόνου. Τέλος, και στις δύο περιπτώσεις ο αριθμός των πελατών που χάνεται είναι σχεδόν ο ίδιος αλλά η διαφορά σε αυτό το κομμάτι θα ήταν πιο αισθητή αν είχαμε περισσότερους γιατρούς.
Η παραπάνω εργασία και κυρίως οι δύο εφαρμογές αποδεικνύονται πολύ χρήσιμες και η συνεισφορά τους είναι πολύτιμη αν σκεφτεί κανείς ότι το ενδιαφέρον του επιχειρηματικού και επιστημονικού κόσμου, αλλά ακόμα και των ανθρώπων στην καθημερινότητά τους, είναι η μείωση της αναμονής σε συστήματα που περιέχουν ουρές εξυπηρέτησης και η ευχαρίστηση των πελατών ώστε να μην εγκαταλείπουν το σύστημα.

Συνοψίζοντας, σαν περαιτέρω έρευνα προτείνεται η μελέτη για την εύρεση τρόπων και μεθόδων που θα μας εξασφαλίσουν τη μείωση της μεταβλητότητας, ώστε να εξασφαλίζονται μικρότεροι χρόνοι αναμονής αλλά και μικρότερες πιθανότητες να είναι όλοι οι εξυπηρετητές απασχολημένοι. Αυτό θα μπορούσε να επιτευχθεί με την μελέτη πολυπλοκότερων συστημάτων και σε άλλες υπηρεσίες, όπως αεροδρόμια, τράπεζες κλπ, και με την προσομοίωσή τους και με διάφορα άλλα πακέτα εκτός του Excel, ελέγχοντας όμως πρώτα αν το πακέτο ταιριάζει με τη λογική του συστήματος. Τέλος, θα πρέπει ο χρήστης να προσέχει ώστε το μοντέλο που χρησιμοποιεί να είναι μία ακριβής αναπαράσταση του συστήματος.

Βιβλιογραφία

Ξένη

- «Queueing Theory», Ivo Adan and Jacques Resing, Department of Mathematics and Computing Science Eindhoven University of Technology.
- J. Medhi 2002, «Stochastic Models in Queueing Theory (Second edition) »
- Kendal D.G 1953, « Stochastic processes occurring in the theory of queues and their analysis by the method of imbedded Markov Chains, Annals of Mathematical Statistics»
- Erlang Agner Krarup, (1909), The Theory of Probabilities and Telephone Conversations
- Andreas Willing 1999, «A short Introduction to Queueing Theory»
- Little John Dutton Conant 1961, «A Proof for the Queuing Formula: L=λW», Operations Research: A Journal of the Institute for Operations Research and the Management Sciences, 9 pp.383-387
- Cachon- Terwiesch 2013, «Matching Supply with Demand: An Introduction to Operations Management» (3rd Edition)
- Samuel Fomundam, Jeffrey Herrmann 2007, «A Survey of Queuing Theory Applications in Healthcare», The Institute for Systems Research
- Linda Green, «QUEUEING THEORY AND MODELING»

Ελληνική

- Νικολαΐδης Χρήστος 2005, «Στοιχεία Πιθανοτήτων και Θεωρίας Ουρών», ΤΕΙ Λάρισας, Τμήμα Τεχνολογίας Πληροφορικής και Τηλεπικοινωνιών
- Βασιλείου Π.– Χ. Γ. 2000, «Στοχαστικές Μέθοδοι στις Επιχειρησιακές Έρευνες», Εκδόσεις Ζήτη
- Φ. Κολυβά – Μαχαίρα, Ε. Μπόρα – Σέντα 1998, «Στατιστική, Θεωρία και Εφαρμογές», Εκδόσεις Ζήτη
- Οικονόμου Γ. Σ και Γεωργίου Α. Κ, 2011, «Ποσοτική ανάλυση για τη λήψη Διοικητικών Αποφάσεων», τόμος Β, Εκδόσεις Μπένου, Αθήνα
- Νικήτας Α. Ασημακόπουλος 1991, «Επιχειρησιακή Έρευνα: Μεθοδολογία, Ουρές Αναμονής, Προσομοίωση», Εκδόσεις Α. Σταμούλης

Ηλεκτρονικές Διευθύνσεις

- www.wikipedia.org
- http://www.kat-hosp.gr/