

ΣΥΣΤΗΜΑΤΑ ΑΝΑΜΟΝΗΣ

Queuing Systems

Μοντέλα Ουρών Markov και Εφαρμογές:

Ουρά M/M/2

Σύστημα M/M/N/K, Erlang-C

Σύστημα M/M/c/c , Erlang-B

Ανάλυση & Σχεδιασμός Τηλεφωνικών Κέντρων

Βασίλης Μάγκλαρης

maglaris@netmode.ntua.gr

18/4/2018

ΠΡΟΣΟΜΟΙΩΣΗ ΣΤΟΧΑΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

BIRTH-DEATH ΟΜΟΙΟΓΕΝΩΝ ΑΦΙΞΕΩΝ *επανάληψη*

- Σε στοχαστικό σύστημα **Birth-Death** με αφίξεις (γεννήσεις) σταθερού μέσου ρυθμού λ ανεξάρτητου του πληθυσμού $n(t) = k$ (**ομοιογενείς αφίξεις Poisson** με ρυθμό $\lambda_k = \lambda$ αφίξεις/sec) οι εργοδικές πιθανότητες (αν υπάρχουν) μπορούν να υπολογισθούν σαν λόγος αφίξεων που βρίσκουν το σύστημα στη κατάσταση $n(t) = k$ στη χρονική διάρκεια T_k , προς τον συνολικό αριθμό αφίξεων σε μεγάλο χρονικό διάστημα παρατήρησης T μιας χρονικής εξέλιξης της διαδικασίας:

$$P_n = \lim_{T \rightarrow \infty} \frac{T_k}{T} = \lim_{T \rightarrow \infty} \frac{\lambda T_k}{\lambda T} \approx \frac{\#\{\text{ΑΦΙΞΕΩΝ στη } n(t) = k \text{ σε ΧΡΟΝΙΚΟ ΔΙΑΣΤΗΜΑ } T_k\}}{\#\{\text{ΣΥΝΟΛΟΥ ΑΦΙΞΕΩΝ σε ΧΡΟΝΙΚΟ ΔΙΑΣΤΗΜΑ } T\}}$$

Άρα μπορούμε να προσομοιώσουμε σύστημα Birth-Death με ομοιογενείς αφίξεις καταμετρώντας τις αφίξεις στις διάφορες καταστάσεις που μεταβαίνει

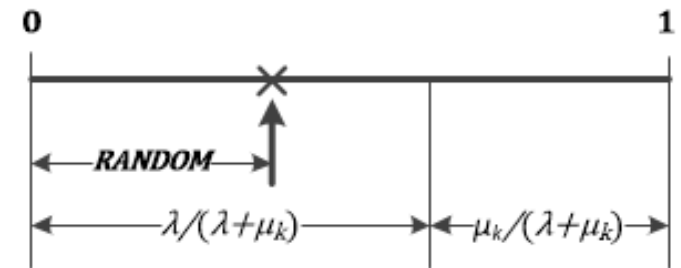
- Η εξέλιξη της κατάστασης (πληθυσμού) του συστήματος προκύπτει από τις πιθανότητες μετάβασης από την κατάσταση $n(t) = k$ στις $(k + 1)$, $(k - 1)$ με το δεδομένο ότι μια από τις δύο μεταβάσεις θα συμβεί με απόλυτη βεβαιότητα:

$$P[k \rightarrow (k + 1)/\text{μετάβαση}] = \lambda/(\lambda + \mu_k), \quad P[k \rightarrow (k - 1)/\text{μετάβαση}] = \mu_k/(\lambda + \mu_k)$$

- Η προσομοίωση ενεργοποιεί τις μεταβάσεις με κλήση τυχαίου αριθμού $RANDOM(0,1)$ ομοιόμορφα κατανεμημένου μεταξύ $(0, 1)$:

$$0 \leq RANDOM(0,1) \leq \frac{\lambda}{\lambda + \mu_k} \Rightarrow \text{ΑΦΙΞΗ}, n(t) \rightarrow k + 1$$

$$\frac{\lambda}{\lambda + \mu_k} < RANDOM(0,1) \leq 1 \Rightarrow \text{ΑΝΑΧΩΡΗΣΗ}, n(t) \rightarrow k - 1$$



- Αν το σύστημα έχει μηδενικό πληθυσμό $n(t) = 0$, η επόμενη μετάβαση είναι πάντα ΑΦΙΞΗ και $n(t) \rightarrow 1$
- Αν το σύστημα δεν επιδέχεται αύξηση πληθυσμού, η $n(t) = K$ είναι blocking state και δεν ενεργοποιείται μετάβαση κατάστασης, αλλά η διαδικασία τυχαίας δημιουργίας επόμενου γεγονότος (ΑΦΙΞΗ ή ΑΝΑΧΩΡΗΣΗ) καθώς και η μέτρηση αφίξεων συνεχίζονται κανονικά

ΠΡΟΣΟΜΟΙΩΣΗ ΣΤΟΧΑΣΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

BIRTH-DEATH ΟΜΟΙΟΓΕΝΩΝ ΑΦΙΞΕΩΝ *επανάληψη*

Στατιστική Σύγκλιση Προσομοίωσης

- Η σύγκλιση της προσομοίωσης ελέγχεται ως προς την **στατιστική σύγκλιση** μεγεθών ενδιαφέροντος, π.χ. εκτίμηση μέσου πληθυσμού $E[n(t)] = \sum_{k=0}^K P_k$ που υπολογίζεται σε τακτά διαστήματα από την αρχή της προσομοίωσης π.χ. κάθε 1000 αφίξεις: (από 0 έως 1000 αφίξεις), (από 0 έως 2000 αφίξεις), (από 0 έως 3000 αφίξεις) κλπ.
- Λόγω **εργοδικότητας** δεν απαιτούνται επαναλήψεις: Η εκτέλεση του προγράμματος και οι μετρήσεις συνεχίζονται σε μία υλοποίηση που διακόπτεται προσωρινά π.χ. κάθε 1000 αφίξεις μέχρι την ικανοποίηση κριτηρίων σύγκλισης
- Η στατιστική σύγκλιση επιταχύνεται αν αγνοήσουμε στη καταμέτρηση αφίξεων στις διάφορες καταστάσεις τις πρώτες μεταβάσεις (π.χ. 1-1000 αφίξεις) που αντιστοιχούν στο **μεταβατικό φαινόμενο** προς την εργοδική κατάσταση
- **Εμπειρικός κανόνας**: Η ταχύτητα σύγκλισης είναι αντιστρόφως ανάλογη με τον βαθμό χρησιμοποίησης του συστήματος
- **Γενική Παρατήρηση**: Η σύγκλιση μιας προσομοίωσης είναι σύνηθες να διερευνάται με μαθηματικά εργαλεία στατιστικής (π.χ. τεστ χ^2) γιατί μπορεί να κρύβονται εξαρτήσεις (correlations) μεγεθών, περιοδικές συμπεριφορές που οδηγούν σε πρόωρους τερματισμούς κλπ. Μια πλήρης προσομοίωση περιλαμβάνει διαστήματα εμπιστοσύνης για τις εκτιμήσεις πιθανοτήτων και ροπών τυχαίων μεταβλητών (π.χ. μέσοι όροι, διασπορά)

ΠΡΟΣΟΜΟΙΩΣΗ ΟΥΡΑΣ Μ/Μ/1/10 επανάληψη

<http://www.netmode.ntua.gr/courses/undergraduate/queues/simulation/>

RANDOM: Ομοιόμορφος τυχαίος αριθμός (0,1)

THRESHOLD: $\lambda / (\lambda + \mu)$

ARRIVALS: Συνολικός αριθμός αφίξεων

ARRIVAL[STATE]: Αριθμός αφίξεων στην κατάσταση STATE = 0, 1, ... , 10

COUNT: Αριθμός μεταβάσεων COUNT = 0, 1, ... , MAXIMUM

STATE: Κατάσταση ουράς (πληθυσμός συστήματος Μ/Μ/1/10) , STATE = 0, 1, ... , 10

P[STATE]: Εργοδική πιθανότητα STATE = 0, 1, ... , 10

AVERAGE: Μέσος πληθυσμός συστήματος Μ/Μ/1/10

INITIALIZE: COUNT = 0, STATE = 0, ARRIVALS = 0, ARRIVAL[0...10] = 0, P[0...10] = 0

ARRIVAL: ARRIVALS = ARRIVALS + 1
ARRIVAL[STATE] = ARRIVAL[STATE] + 1
COUNT = COUNT + 1
IF STATE = 10 : **GO TO** LOOP
ELSE : STATE = STATE + 1
GO TO LOOP

LOOP : **IF** STATE = 0 : **GO TO** ARRIVAL
ELSE :
IF RANDOM < THRESHOLD : **GO TO** ARRIVAL
ELSE : **GO TO** DEPARTURE

DEPARTURE : COUNT = COUNT + 1 ; STATE = STATE - 1
IF COUNT < MAXIMUM : **GO TO** LOOP
ELSE : P[STATE=1...10] = ARRIVAL[STATE= 1...10] / ARRIVALS
AVERAGE = SUM { STATE ^ P[STATE] }, STATE = [1...10]

ΠΡΟΣΟΜΟΙΩΣΗ ΟΥΡΑΣ Μ/Μ/1/10

Εργαστηριακή Άσκηση

Οδηγίες Προγράμματος Προσομοίωσης Ουράς Μ/Μ/1/10

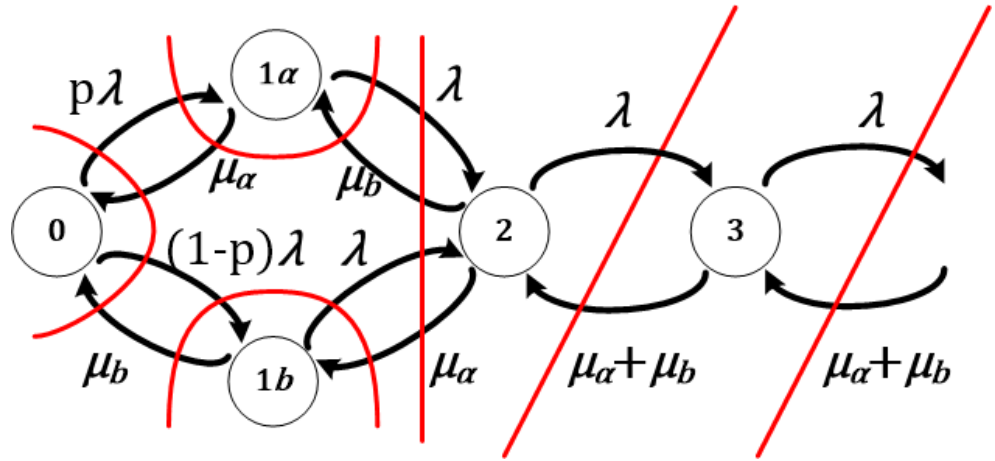
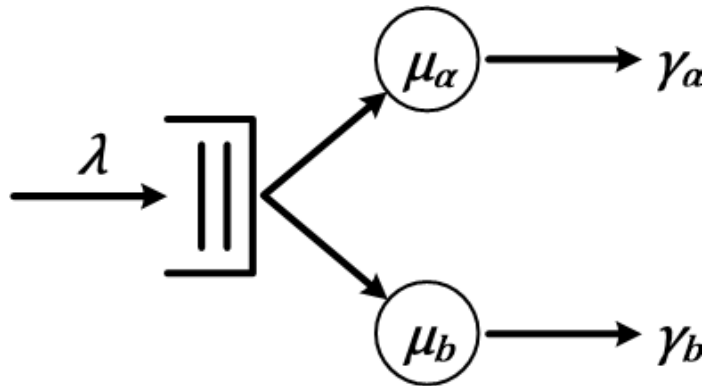
- Δημιουργία προγράμματος με βάση τον ψευτο-κώδικα της προηγούμενης διαφάνειας
- Γλώσσα προγραμματισμού: Γενικής φύσεως (Octave/MATLAB, Python) και όχι εξειδικευμένη γλώσσα Προσομοίωσης
- Υπολογισμός εργοδικών πιθανοτήτων P_k , $E(k) = \sum_{k=0}^{10} P_k$ και $P_{\text{blocking}} = P_{10}$ για $\lambda = 1$ και $1/\mu = 0.5, 0.75, 0.85$ με μετρητές αφίξεων στις καταστάσεις $n(t) = k$ από την αρχή μέχρι 5.000, 10.000, 15.000 ... αφίξεις. Σύγκριση με γνωστά αναλυτικά αποτελέσματα ουράς Μ/Μ/1/10:

$$P_k = P_0 \left(\frac{\lambda}{\mu} \right)^k, \quad \sum_{k=0}^{10} P_k = 1$$

- Μέτρο σύγκλισης: Διαφορά μεταξύ διαδοχικών μετρήσεων του μέσου μήκους ουράς (από την αρχή μέχρι 5.000, 10.000, 15.000... αφίξεις) μέχρι εξαντλήσεως του διαθέσιμου χρόνου αν δεν διαπιστωθεί σύγκλιση νωρίτερα (όριο 200.000 γεγονότα, αφίξεις + εξυπηρετήσεις)
- **Υπενθύμιση:** Η στατιστική σύγκλιση επιταχύνεται αν αγνοήσουμε στη καταμέτρηση αφίξεων στις διάφορες καταστάσεις τις πρώτες μεταβάσεις (π.χ. 1-5.000 αφίξεις) που αντιστοιχούν στο **μεταβατικό φαινόμενο** προς την εργοδική κατάσταση
- **Γενική Οδηγία:** Για debugging εξάγεται αρχικά λεπτομερές **trace** των μεταβάσεων της κατάστασης του συστήματος το οποίο θα απενεργοποιηθεί μετά την επιβεβαίωση της ορθότητας του κώδικα

ΟΥΡΑ Μ/Μ/2

- Αφίξεις Poisson με ομοιόμορφο μέσο ρυθμό $\lambda_k = \lambda$
- 2 ανεξάρτητοι εκθετικοί εξυπηρετητές (α), (b) με άνισους ρυθμούς μ_α και μ_b
- Άπειρη Χωρητικότητα
- Άφιξη σε άδειο σύστημα δρομολογείται στον (α) με πιθανότητα p και στον (b) με πιθανότητα $(1 - p)$



Εξισώσεις Ισορροπίας:

$$\lambda P_0 = \mu_\alpha P_{1\alpha} + \mu_b P_{1b}$$

$$(\lambda + \mu_\alpha) P_{1\alpha} = p\lambda P_0 + \mu_b P_2$$

$$(\lambda + \mu_b) P_{1b} = (1-p)\lambda P_0 + \mu_\alpha P_2$$

$$\lambda(P_{1\alpha} + P_{1b}) = (\mu_\alpha + \mu_b) P_2, \quad \lambda P_k = (\mu_\alpha + \mu_b) P_{k+1}, \quad k = 2, 3, \dots$$

$$P_0 + P_{1\alpha} + P_{1b} + P_2 + P_3 + \dots = 1, \quad \lambda/(\mu_\alpha + \mu_b) < 1 \text{ για σύγκληση (εργοδικότητα)}$$

Βαθμοί Χρησιμοποίησης – Ρυθμαποδόσεις Εξυπηρετητών:

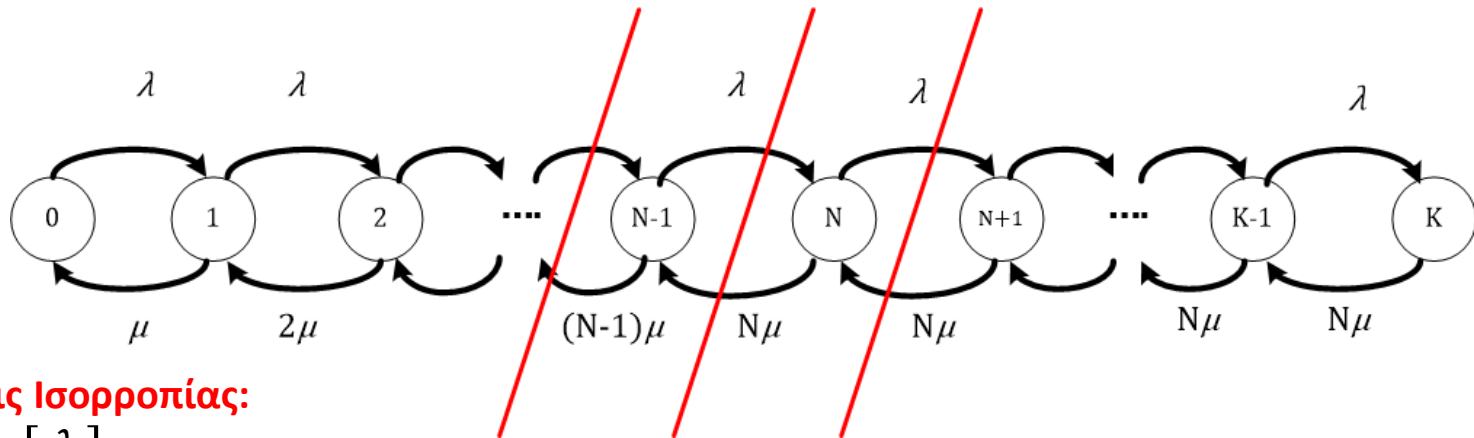
$$U_\alpha = 1 - P_0 - P_{1b} \quad \gamma_\alpha = \mu_\alpha U_\alpha$$

$$U_b = 1 - P_0 - P_{1\alpha} \quad \gamma_b = \mu_b U_b$$

$$\gamma = \lambda = \gamma_\alpha + \gamma_b$$

ΟΥΡΑ Μ/Μ/Ν/Κ

- Αφίξεις Poisson με ομοιόμορφο μέσο ρυθμό $\lambda_k = \lambda$
- Ν ανεξάρτητοι εκθετικοί εξυπηρετητές με ίσους ρυθμούς μ
- Χωρητικότητα Κ, $N \leq K$ (π.χ. **call center** με Ν εξυπηρετητές & δυνατότητα αναμονής μέχρι $K - N$ κλήσεις)
- Μέσοι ρυθμοί εξυπηρέτησης στη κατάσταση k :
 - $\mu_k = k\mu$, $k = 1, 2, \dots, N$
 - $\mu_k = N\mu$, $k = N, N + 1, \dots, K - 1, K$
- Εργοδική κατάσταση $n(t)$: Αριθμός πελατών στο σύστημα, αδιάφορα από χρήση συγκεκριμένων εξυπηρετητών (π.χ. σε σύστημα Μ/Μ/2, $\mu_a = \mu_b = \mu$, $P_1 = P_{1a} + P_{1b}$)



Εξισώσεις Ισορροπίας:

$$P_k = \left[\frac{\lambda}{k\mu} \right] P_{k-1}, \quad k = 1, 2, \dots, N - 1$$

$$P_k = \left[\frac{\lambda}{N\mu} \right] P_{k-1}, \quad k = N, N + 1, \dots, K - 1, K$$

$$P_0 + P_1 + \dots + P_{K-1} + P_K = 1, \quad P_K = P_{\text{blocking}}, \quad \gamma = \lambda (1 - P_{\text{blocking}})$$

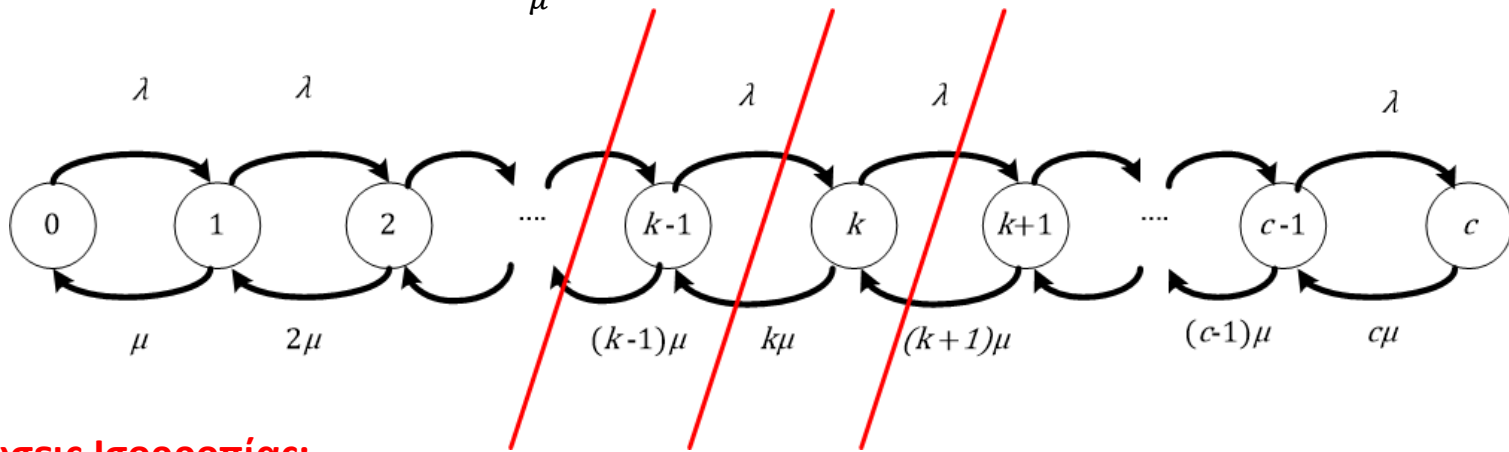
$$P_{\text{waiting}} = P_N + P_{N+1} + \dots + P_K = 1 - (P_0 + P_1 + \dots + P_{N-1}) \quad (\text{Erlang-C Formula})$$

ΟΥΡΑ Μ/Μ/Ϸ/Ϸ

(τηλεφωνικό κέντρο με Ϸ εξωτερικές γραμμές, trunks)

- Αφίξεις Poisson με ομοιόμορφο μέσο ρυθμό $\lambda_n = \lambda$ (εξωτερικές κλήσεις - τηλεφωνήματα/sec)
- Ϸ ανεξάρτητοι εκθετικοί εξυπηρετητές (εξωτερικές γραμμές τηλεφωνικού κέντρου)
- Χωρητικότητα Ϸ πελάτες (τηλεφωνήματα, εξωτερικές κλήσεις)
- Ρυθμοί εξυπηρέτησης στη κατάσταση k :

$$\mu_k = k\mu, \quad k = 1, 2, \dots, \quad \frac{1}{\mu} = E(s): \text{ Μέση Διάρκεια Τηλεφωνήματος (π.χ. 3 min ή 180 sec)}$$



Εξισώσεις Ισορροπίας:

$$P_k = \left[\frac{\lambda}{k\mu} \right] P_{k-1} = \left(\frac{\rho^k}{k!} \right) P_0, \quad k = 1, 2, \dots, \quad c \quad \rho \triangleq \frac{\lambda}{\mu} \text{ Erlangs}$$

$$P_0 + P_1 + \dots + P_{c-1} + P_c = 1 \Rightarrow P_0 = \frac{1}{\sum_{k=0}^c \frac{\rho^k}{k!}}$$

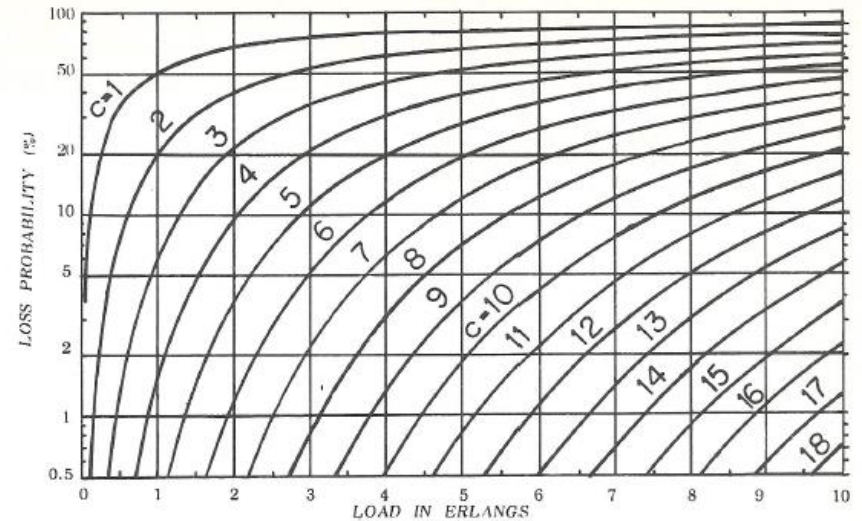
$$P_c = P_{\text{blocking}} = \frac{\rho^c / c!}{\sum_{k=0}^c \frac{\rho^k}{k!}} \triangleq B(\rho, c) \text{ (Erlang-B Formula)}$$

ΠΙΝΑΚΕΣ Erlang B(ρ, c)

Αναδρομικός Υπολογισμός

$$B(\rho, 0) = 1$$

$$B(\rho, n) = \frac{\rho B(\rho, n-1)}{\rho B(\rho, n-1) + n}, \quad n = 1, 2, \dots, c$$



ρ	c-1	2	3	4	5	6	7	8
0.00	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.25	.2000	.0244	.0020	.0001	**	*	*	*
0.50	.3333	.0769	.0127	.0016	.0002	**	*	*
0.75	.4286	.1385	.0335	.0062	.0010	.0001	**	*
1.00	.5000	.2000	.0625	.0154	.0031	.0005	.0001	**
1.25	.5556	.2577	.0970	.0294	.0073	.0015	.0003	**
1.50	.6000	.3103	.1343	.0480	.0142	.0035	.0008	.0001
1.75	.6364	.3577	.1726	.0702	.0240	.0070	.0017	.0004
2.00	.6667	.4000	.2105	.0952	.0367	.0121	.0034	.0009
2.25	.6923	.4378	.2472	.1221	.0521	.0192	.0061	.0017
2.50	.7143	.4717	.2822	.1499	.0697	.0282	.0100	.0031
2.75	.7333	.5021	.3152	.1781	.0892	.0393	.0152	.0052
3.00	.7500	.5294	.3462	.2062	.1101	.0522	.0219	.0081
3.25	.7647	.5541	.3751	.2336	.1318	.0666	.0300	.0121
3.50	.7778	.5765	.4021	.2603	.1541	.0825	.0396	.0170
3.75	.7895	.5968	.4273	.2860	.1766	.0994	.0506	.0232
4.00	.8000	.6154	.4507	.3107	.1991	.1172	.0628	.0304
4.25	.8095	.6324	.4725	.3343	.2213	.1355	.0760	.0388
4.50	.8182	.6480	.4929	.3567	.2430	.1542	.0902	.0483
4.75	.8261	.6624	.5119	.3781	.2643	.1730	.1051	.0587
5.00	.8333	.6757	.5287	.3983	.2849	.1919	.1205	.0701
5.25	.8400	.6880	.5463	.4176	.3048	.2106	.1364	.0822
5.50	.8462	.6994	.5618	.4358	.3241	.2290	.1525	.0949
5.75	.8519	.7101	.5764	.4531	.3426	.2472	.1688	.1082
6.00	.8571	.7200	.5902	.4696	.3604	.2649	.1851	.1219
6.25	.8621	.7293	.6031	.4852	.3775	.2822	.2013	.1359
6.50	.8667	.7380	.6152	.4999	.3939	.2991	.2174	.1501
6.75	.8710	.7462	.6267	.5140	.4096	.3155	.2333	.1644
7.00	.8750	.7539	.6376	.5273	.4247	.3313	.2489	.1788
7.25	.8788	.7611	.6478	.5401	.4392	.3467	.2642	.1932

ρ	c-9	10	11	12	13	14	15	16
0.00	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
0.25								
0.50								
0.75								
1.00	*							
1.25	**							
1.50	**	*						
1.75	.0001	**	*					
2.00	.0002	**	**					
2.25	.0004	.0001	**	*				
2.50	.0009	.0002	.0001	**	*			
2.75	.0016	.0004	.0001	**	**			
3.00	.0027	.0008	.0002	.0001	**	*		
3.25	.0043	.0014	.0004	.0001	**	**		
3.50	.0066	.0023	.0007	.0002	.0001	**	*	
3.75	.0096	.0036	.0012	.0004	.0001	**	**	
4.00	.0133	.0053	.0019	.0006	.0002	.0001	**	*
4.25	.0180	.0076	.0029	.0010	.0003	.0001	**	**
4.50	.0236	.0105	.0043	.0016	.0006	.0002	.0001	**
4.75	.0301	.0141	.0060	.0024	.0009	.0003	.0001	**
5.00	.0375	.0184	.0083	.0034	.0013	.0005	.0002	.0001
5.25	.0457	.0234	.0111	.0048	.0019	.0007	.0003	.0001
5.50	.0548	.0293	.0144	.0066	.0028	.0011	.0004	.0001
5.75	.0647	.0358	.0184	.0087	.0039	.0016	.0006	.0002
6.00	.0751	.0431	.0230	.0114	.0052	.0022	.0009	.0003
6.25	.0862	.0511	.0282	.0145	.0069	.0031	.0013	.0005
6.50	.0978	.0598	.0341	.0181	.0090	.0042	.0018	.0007
6.75	.1098	.0690	.0406	.0223	.0115	.0055	.0025	.0010
7.00	.1221	.0787	.0477	.0271	.0144	.0071	.0033	.0015
7.25	.1347	.0890	.0554	.0324	.0177	.0091	.0044	.0020

ΠΑΡΑΔΕΙΓΜΑ ΑΝΑΛΥΣΗΣ ΤΗΛΕΦΩΝΙΚΟΥ ΚΕΝΤΡΟΥ

- Τηλεφωνικό Κέντρο με 7 εξωτερικές γραμμές προωθεί κίνηση (προς τις 2 κατευθύνσεις) με μέσο ρυθμό κλήσεων 2 κλήσεις το λεπτό με μέση διάρκεια κλήσης 3 min
- Θεωρώ ότι οι εξωτερικές κλήσεις ακολουθούν διαδικασία **Poisson** με μέσο ρυθμό $\lambda = 2$ κλήσεις/min και χρόνο εξυπηρέτησης **Εκθετικό** με μέση διάρκεια $1/\mu = 3$ min, άρα το συνολικό προσφερόμενο φορτίο (**offered traffic**) είναι

$$\rho = \lambda/\mu = 6 \text{ Erlangs}$$

- Υποθέτουμε πως οι κλήσεις που δεν βρίσκουν γραμμή χάνονται οριστικά. Άρα η πιθανότητα απώλειας δίνεται από τον τύπο

$$B(\rho, c) = B(6,7) = 18.51\%$$

- Το εξυπηρετούμενο φορτίο (**carried traffic**) είναι

$$\rho[1 - B(\rho, c)] = \frac{\lambda}{\mu} [1 - B(\rho, c)] = \frac{\gamma}{\mu} = 4.8894 \text{ Erlangs}$$

- Το φορτίο υπερχείλισης (**overflow traffic**) είναι

$$\rho B(\rho, c) = 1.1106 \text{ Erlangs}$$

ΠΑΡΑΔΕΙΓΜΑ ΣΧΕΔΙΑΣΜΟΥ ΤΗΛΕΦΩΝΙΚΟΥ ΚΕΝΤΡΟΥ

- Τηλεφωνικό Κέντρο με c εξωτερικές γραμμές (trunks) προωθεί κίνηση (προς τις 2 κατευθύνσεις) με μέσο ρυθμό κλήσεων 2 κλήσεις το λεπτό με μέση διάρκεια κλήσης 3 min.
- Θεωρώ ότι οι εξωτερικές κλήσεις ακολουθούν διαδικασία **Poisson** με μέσο ρυθμό $\lambda = 2$ κλήσεις/min και χρόνο εξυπηρέτησης **Εκθετικό** με μέση διάρκεια $\frac{1}{\mu} = 3$ min, άρα το συνολικό προσφερόμενο φορτίο (**offered traffic**) είναι

$$\rho = \frac{\lambda}{\mu} = 6 \text{ Erlangs}$$

- Υποθέτουμε πως οι κλήσεις που δεν βρίσκουν γραμμή χάνονται οριστικά. Ζητείται ο απαιτούμενος αριθμός εξωτερικών γραμμών (trunks) c ώστε ο ρυθμός απωλειών (**Grade of Service, GOS**) να είναι μικρότερος από 0.3%
- Από τους πίνακες προκύπτει πως $B(6,13) = 0.52\%$ και $B(6,14) = 0.24\%$, άρα οι απαιτήσεις καλύπτονται με ελάχιστο αριθμό εξωτερικών γραμμών $c = 14$ trunks