

Uma análise crítica da Lei Municipal 13.948 ou “Lei das Filas” sob a ótica da Pesquisa Operacional: conclusões derivadas de modelos de simulação de eventos discretos

Leonardo Chwif (UNIFEO) leonardochwif@unifio.br
Afonso Celso Medina (USP) afmedina@uol.com.br

Resumo

Este trabalho discute, sob a ótica da Pesquisa Operacional, a aplicabilidade da lei municipal 13.948 da cidade de São Paulo-SP, conhecida como “lei das filas”, que obriga as agências bancárias a atenderem os clientes em um tempo de fila inferior a 15 minutos. Três aspectos são considerados: a adequabilidade do critério do tempo de fila como regra de atendimento, o impacto sobre as chamadas “filas preferenciais” e o valor da multa como um estímulo ou não ao cumprimento da lei. Três modelos de simulação foram construídos e os resultados apresentados corroboram a validade do critério de 15 minutos.

Palavras-chave: Filas em bancos, Simulação.

1. Introdução

Em 2005 entrou em vigor, no município de São Paulo, a Lei Municipal 13.948 (SÃO PAULO, 2005b) que determina no seu Art. 1º: *Ficam as agências bancárias e demais estabelecimentos de crédito do Município de São Paulo obrigados a colocar à disposição dos usuários pessoal suficiente no setor de caixas, para que o atendimento seja feito em prazo hábil, respeitada a dignidade e o tempo do usuário.*

No seu Art. 2º, a lei determina que o “prazo hábil” do artigo anterior é de 15 minutos em dias normais, 25 minutos às vésperas e depois de feriados prolongados e 30 minutos nos dias de pagamentos de funcionários públicos, não podendo ultrapassar tais prazos. Em função do critério de prazo adotado, a lei é hoje conhecida entre a população como “lei das filas”. Em outros estados, leis semelhantes têm sido adotadas, sempre com foco na regulamentação das filas em agências bancárias a partir do critério do “tempo máximo de atendimento”.

A repercussão da lei foi enorme (BALBI, 2005), motivando, até mesmo, um Mandado de Segurança por parte da Federação Brasileira de Bancos – FEBRABAN que suspendeu os seus efeitos a partir do dia 9 de maio de 2006.

O objetivo deste trabalho é analisar esta lei sob a ótica da Pesquisa Operacional, de modo a responder às seguintes indagações:

- a) O indicador de 15 minutos é adequado como critério de atendimento?
- b) Qual o seu impacto sobre o sistema de filas múltiplas e de filas preferenciais?
- c) O valor da multa garante o atendimento em 15 minutos ou é um estímulo ao não cumprimento da lei?

A continuação deste trabalho está organizada como descrita a seguir. A seção 2 discute a “lei das filas”, sua repercussão e reconstrói seu significado à luz do problema de otimização envolvido. A seção 3 descreve a metodologia utilizada e os modelos de simulação construídos. A seção 4 apresenta os resultados computacionais obtidos e, finalizando, a seção 5 é reservada para conclusões e recomendações para trabalhos futuros.

2. A “lei das filas” e seu problema de otimização

À época da entrada em vigor da citada lei, foram publicados dois artigos no jornal *Folha de S. Paulo*: um (CYPRIANO, 2005) em que a lei é discutida sob a perspectiva da Federação Brasileira dos Bancos – FEBRABAN e outro (MORAES, 2005) em que ela é discutida sob a perspectiva da Confederação Nacional dos Bancários – CNB. Cypriano (2005) argumenta que a lei “não resolve o problema” e que o problema está na “concentração de pagamentos e de recebimentos nos cinco primeiros dias úteis de todo mês”. O artigo aponta também que o caminho de racionalização das filas adotado pelo sistema bancário é o da redução do tempo de atendimento, seja por um atendimento mais ágil no caixa, seja pelo investimento em tecnologia de informação, visível no maior número de caixas eletrônicos ou na oferta de sistemas do tipo *homebanking*. Moraes (2005), por outro lado, argumenta que o problema está no “número menor de empregados do que o necessário (basta ver o número de caixas fechados nas agências)”. Apesar de intrinsecamente ligadas aos interesses patronais e sindicais, as duas argumentações refletem os dois modos básicos de redução do tamanho de uma fila: ou pela diminuição do tempo de atendimento – como sugere Cypriano (2005) – ou pelo aumento do número de servidores – como sugere Moraes (2005).

Ressalvem-se alguns aspectos nos artigos supracitados. É evidente que a lei “não resolve o problema”, pois se trata tão somente do poder público intervindo como agente regulamentador do sistema, de modo a garantir que o atendimento “seja feito em prazo hábil, respeitada a dignidade e o tempo do usuário” (SÃO PAULO, 2005b). A alegada excepcionalidade do surgimento das filas nos primeiros dias do mês está considerada no Art. 2º da lei, que altera o critério de atendimento para 30 minutos nesses dias. Portanto, a lei é um passo na garantia de nível de serviço mínimo ao usuário e já prevê os estados de congestionamento do sistema. Com relação a simplesmente aumentar o número de caixas, que ao senso comum é a melhor solução; não é, à luz da Pesquisa Operacional, a maneira correta de se identificar a solução ótima do problema. O que a Pesquisa Operacional pretende é fornecer aos administradores subsídios para suas decisões, baseado na construção de modelos científicos do sistema em estudo, incorporando fatores probabilísticos, com os quais é possível prever e comparar o resultado de alternativas de decisão.

A questão que inicialmente deve ser respondida é: o critério de tempo de atendimento é adequado ou não como indicador do nível de serviço? O decreto municipal 45.939 (SÃO PAULO, 2005a), que regulamenta a lei, estabelece que “tempo de atendimento” é o tempo computado a partir do ingresso do usuário “na fila de atendimento do setor de caixas, encerrando-se no momento em que se iniciar seu atendimento”. Não se considera, portanto, o tempo de serviço no caixa, que também é uma parcela importante do tempo em que o usuário permanece na agência bancária. De fato, o tempo de fila é aquele que o senso comum considera como mais importante, ou pelo menos, é o tempo ao qual a maioria dos usuários parece ser mais sensível (LARSON, 1987). Esse critério, portanto, atende de imediato aos interesses da maioria da população por um limite máximo de espera em fila.

Estabelecida a regra de atendimento, o próximo passo é estudar a sua aplicação. Tem-se, do ponto de vista operacional, um problema de dimensionamento de filas que se desdobra em um problema maior de otimização: qual é o número de caixas adequado para atendimento da lei ao menor custo total? Classicamente, esse é um problema da “teoria das filas” (KLEINROCK, 1975). A estimativa usual do custo neste tipo de problema considera o custo dos servidores (ou caixas bancários) e o custo de espera do usuário. O custo dos servidores pode ser obtido diretamente a partir dos custos de encargos e salários dos operadores de caixas, enquanto o custo de espera do usuário é de estimativa mais difícil, pois incorpora aspectos não tangíveis

como a insatisfação, desejo de mudança de banco, perdas de negócios futuros etc. Por outro lado, a lei estabelece também uma multa a ser aplicada no caso do seu desrespeito: R\$ 564,00, dobrada no caso de reincidência. Esta multa pode ser entendida pelo Banco como uma alternativa ao custo de se colocar mais caixas abertos – em uma situação semelhante àquela encontrada em companhias de aviação que preferem operar em *overbooking* e pagar as multas adicionais a operar suas aeronaves com folga de ocupação e reduzir o faturamento dos vôos. Moraes (2005) ilustra essa situação ao afirmar que os bancos “transferem o custo” do sistema para o cliente “que fica esperando”.

Outro aspecto importante diz respeito às chamadas “filas preferenciais”. Elas são um meio de atender idosos, gestantes ou deficientes físicos que, amparados por diversas leis, devem receber atendimento preferencial. Contudo, em algumas agências, é usual a existência de uma terceira fila para “clientes especiais” que, segundo critérios comerciais próprios de cada banco, também são merecedores de um atendimento diferenciado. Quando um sistema deixa de operar em fila única, ele quebra a regra do “primeiro a chegar é o primeiro a ser atendido”, possibilitando o surgimento de injustiças sociais, como bem relata o trabalho de Larson (1987). Clientes que utilizam filas especiais estão, de fato, desrespeitando a ordem de chegada e, na expressão popular, “furando a fila”. Portanto, outra questão que a ser analisada é qual o impacto da lei em sistemas que não possuem uma fila única?

3. Metodologia

Usualmente, duas técnicas da Pesquisa Operacional são utilizadas para o estudo de sistemas de filas: a teoria das filas ou a simulação computacional. A teoria das filas é baseada na construção de modelos analíticos do sistema, enquanto a simulação utiliza-se de um modelo computacional para representar o sistema em estudo.

Neste trabalho optou-se por modelar o problema pela simulação, em detrimento da teoria das filas, pois não é interesse aqui discutir o método de dimensionamento de filas, mas analisar a lei das filas. Além disso, a teoria das filas não possui expressões analíticas prontas para sistemas com fila em paralelo, como é o caso de sistemas com filas preferenciais. O leitor interessado pode consultar o trabalho de Chwif e Medina (2006) que discute casos de pertinência da teoria das filas e dos modelos de simulação.

Foram desenvolvidos 3 modelos de simulação para responder às questões levantadas na seção anterior:

- a) Modelo 1: modelo de fila única (figura 1a)
- b) Modelo 2: modelo de filas múltiplas (figura 1b)
- c) Modelo 3: modelo de fila única com adição de fila preferencial (figura 2)

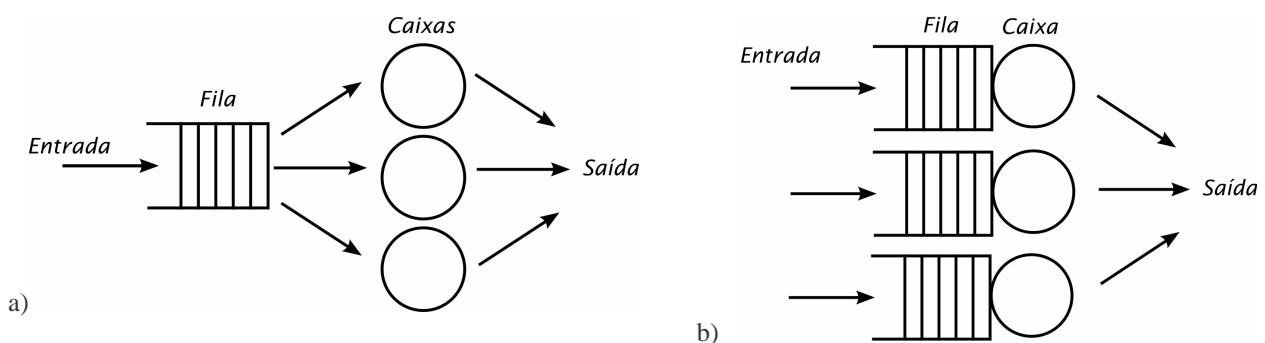


Figura 1 – a) modelo 1, com fila única e vários caixas de atendimento; e b) modelo 2, com filas múltiplas

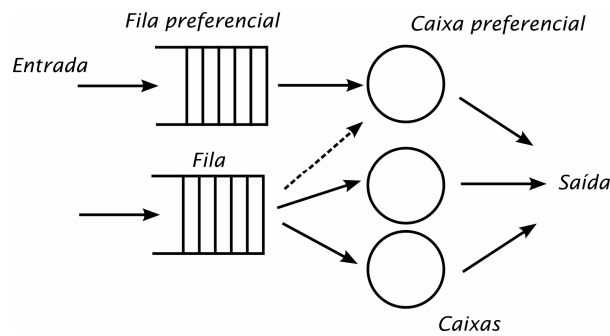


Figura 2 – Modelo 3, fila única e fila preferencial

Os modelos desenvolvidos seguiram a metodologia de simulação (figura 3), que está dividida em 3 fases, a saber: concepção do sistema (compreende a etapa de definição dos objetivos, modelagem abstrata e conceitual), implementação do modelo computacional (em um software de simulação) e análises (obtenção dos resultados experimentais). Os modelos foram implementados no software Simul8 para simulação de eventos discretos da Simul8 Co.

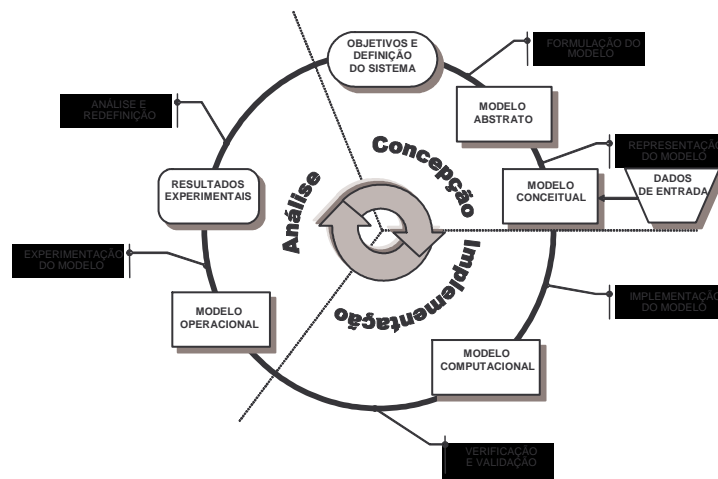


Figura 3 – Metodologia da simulação (CHWIF E MEDINA, 2006)

Para a construção de todos os modelos foram adotadas as seguintes premissas¹:

- Adota-se um fluxo de 100 clientes por hora, segundo uma distribuição de Poisson;
- Os tempos de atendimentos nos caixas seguem uma distribuição triangular com menor valor de 0,5 minutos, moda de 2,5 minutos e menor valor de 10 minutos²;
- O custo do cliente que fica mais de quinze minutos na fila é de R\$ 564,00, relativo à multa estabelecida em lei, e o dobro desse valor na reincidência. Considera-se, portanto, que todo cliente não atendido como determina a lei, formula uma queixa à prefeitura;
- O custo médio mensal de cada caixa, já incluindo benefícios, é de R\$ 2.494,00;
- Almoço, concessões e faltas serão supostas como cobertas por uma equipe extra de caixas e serão desconsiderados do modelo;

¹ Quando não informado o contrário, os dados utilizados foram obtidos a partir de um estudo realizado para uma agência bancária típica na cidade de São Paulo-SP.

² A curva ajustada em software de *fitting* dos tempos de atendimento nos caixas obtida a partir de dados reais, apresentou uma distribuição *lognormal*. Contudo, preferiu-se adotar a distribuição triangular pois a distribuição *lognormal* possui uma cauda para a direita, podendo gerar tempos muito superiores a 10 minutos, o que na prática é muito difícil de ocorrer.

- f) Para o modelo 2 (filas múltiplas) considera-se que o cliente escolhe a fila que tiver o menor número de pessoas;
- g) Para o modelo 3 (atendimento preferencial) adota-se que 10% do fluxo de clientes são idosos, gestantes ou deficientes físicos. Considera-se que, quando não há clientes para o atendimento preferencial, o caixa atenderá um cliente regular que esteja aguardando em fila.

4. Resultados

O modelo 1 foi simulado para diversos números de caixas, obtendo-se a tabela 1, que ilustra para cada número de caixas o custo total (custo de caixas mais custo da multa) e a porcentagem de clientes que esperam menos do que 15 minutos (indicadores principais do modelo). Os resultados apresentados na tabela 1 são as médias de várias replicações.

Número de caixas >	5	6	7	8	9	10
% de clientes atendidos em menos de 15 min	11,20	23,62	75,18	99,92	100,00	100,00
Custo total (R\$/mês)	864.899	653.976	210.684	21.305	22.446	24.940
				(ótimo)		

Tabela 1 – Resultados da simulação do modelo 1

Pela tabela 1, percebe-se que no ponto de menor custo total para a agência, a porcentagem de clientes atendidos em um tempo inferior a 15 minutos é de 99,9%, ou seja, um cliente em cada 1.000 aguardando mais de 15 minutos. O que esse resultado mostra é que o critério de tempo de atendimento, dentro dos valores adotados, está próximo do ponto de custo ótimo da agência. Considerando o custo do caixa, pode-se estimar, para este cenário, qual seria o custo com que o banco “enxerga” o tempo de “espera em fila” do cliente. Este custo representa o custo total do banco dividido pelo tempo total que os clientes esperaram em filas durante o mês (figura 4). Para a solução ótima, 8 caixas, o custo de espera é de R\$ 55,53 por hora que um cliente aguarda em fila. Este valor é, notadamente, um valor bem inferior aos R\$ 8,91 por hora no sistema com um caixa a menos. Cabe destacar que o tempo médio de espera em fila com 8 caixas é de 2,1 minutos e, para 7 caixas, o tempo médio de espera é de 11,5 minutos.

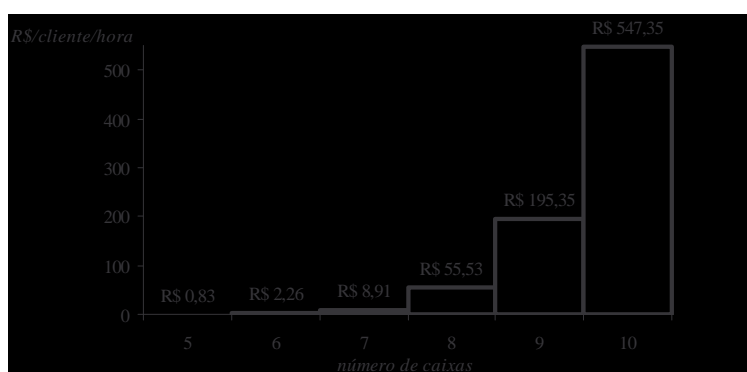


Figura 4 – Custos de espera unitários (por cliente) em função do número de caixas

Analisando-se os resultados do modelo 2 (tabela 2), o menor custo total ocorre com 10 caixas abertos e um nível de atendimento próximo a 100%. Comparando-se esse resultado com o obtido para o modelo 1, pode-se concluir que o desempenho da fila múltipla é ligeiramente pior, pois para um mesmo nível de atendimento, aproximadamente 99,9%, necessita-se de nove caixas.

No modelo 3 existem duas filas de atendimento, uma regular e outra somente para

atendimento preferencial. Os principais resultados estão relacionados na tabela 3. Neste caso, há um desbalanço entre os níveis de serviço das duas filas e, no ponto de custo ótimo para a agência, o percentual de clientes preferenciais atendidos é de 91,82%. Os resultados da tabela 3 demonstram que quando a relação do número de caixas normais e preferenciais está mal dimensionada, o atendimento preferencial pode acabar sendo injusto para uma das partes (normal ou preferencial), pois o nível de serviço dos clientes preferenciais raramente se iguala ao dos clientes regulares. Na realidade, podem ocorrer dois fenômenos: se o número de caixas regulares (não preferenciais) for pequeno em relação aos preferenciais, o atendimento preferencial prejudica o atendimento regular. Outra injustiça ocorre, ironicamente, no outro extremo com os próprios necessitados do atendimento preferencial. Se o número de caixas regulares é alto e se o número de caixas de atendimento preferencial não está bem dimensionado, como é o caso do modelo 3, o atendimento regular terá um melhor nível de serviço do que o atendimento preferencial.

	Número de Caixas >							
	5	6	7	8	9	10	11	
% de clientes atendidos em menos de 15 min na fila 1	10,10	23,62	79,43	99,00	99,96	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 2	12,06	26,57	78,62	98,93	100,00	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 3	12,45	26,29	79,04	99,19	99,96	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 4	11,55	26,67	78,42	98,86	99,87	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 5	12,44	25,88	78,43	99,20	100,00	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 6		25,09	78,04	99,35	99,78	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 7			77,68	99,42	100,00	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 8				99,29	100,00	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila 9					99,92	99,95	99,98	
% de clientes atendidos em menos de 15 min na fila 10						100,00	100,00	
Custo total (R\$/mês)	887.685	692.238	313.106	53.228	25.717	25.052	27.546	(ótimo)

Tabela 2 – Resultados da simulação do modelo 2

	Número de Caixas >							
	5	6	7	8	9	10	11	12
% de clientes atendidos em menos de 15 min na fila REGULAR	10,78	21,90	73,26	99,30	100,00	100,00	100,00	100,00
% de clientes atendidos em menos de 15 min na fila PREFERENCIAL	81,26	85,42	88,94	91,20	91,82	91,89	92,09	92,39
Custo total (R\$/mês)	866.253	663.676	265.732	84.360	80.538	85.062	84.398	86.215
					(ótimo)			

Tabela 3 – Resultados da simulação do modelo 3

A tabela 4 representa a mesma situação do modelo 3, porém, com 2 caixas dedicados aos clientes preferenciais. Neste caso, há um equilíbrio entre os níveis de serviço de ambas as filas e o custo total ótimo é 72% inferior ao custo total ótimo da situação desbalanceada.

	Número de caixas >						
	5	6	7	8	9	10	
% de clientes atendidos em menos de 15 min na fila REGULAR	10.19	21,30	74,06	99,64	100,00	100,00	
% de clientes atendidos em menos de 15 min na fila PREFERENCIAL	99.97	100,00	100,00	100,00	100,00	100,00	
Custo total (R\$/mês)	868.960	660.969	225.122	24.238	22.446	24.940	(ótimo)

Tabela 4 – Resultados do modelo 3, com 2 caixas preferenciais

A tabela 5 apresenta os resultados obtidos para a simulação do sistema representado pelo modelo 3 acrescido de uma fila adicional para clientes especiais. Neste caso, adotou-se que a participação dos clientes especiais, preferenciais e regulares, correspondem a, respectivamente, 20%, 10% e 70% do fluxo total de clientes na agência. Quando se compara os resultados do modelo 3 alterado com os resultados do modelo 1 (fila única), nota-se um aumento de 6% no custo total ótimo. Conclui-se, portanto, que quanto maior for a segregação no atendimento maior é o custo para um mesmo nível de atendimento.

	Número de caixas >					
	5	6	7	8	9	10
% de clientes atendidos em menos de 15 min na fila REGULAR	6.49	17.12	67.42	98.66	99.98	100,00
% de clientes atendidos em menos de 15 min na fila ESPECIAL	100,00	100,00	100,00	100,00	100,00	100,00
% de clientes atendidos em menos de 15 min na fila PREFERENCIAL	100,00	100,00	100,00	100,00	100,00	100,00
Custo total (R\$/mês)	887.910	682.853	284.004	33.375	22.558	24.940
					(ótimo)	

Tabela 5 – Resultados do modelo 3 alterado, com 2 caixas preferenciais e 2 caixas especiais

Para realizar uma análise de sensibilidade quanto ao valor da multa, foram utilizados dois cenários: (1) o valor da multa é a metade do valor de referência e (2) o valor da multa é o dobro do valor de referência. Os resultados estão dispostos respectivamente nas tabelas 6 e 7. Note-se que a sensibilidade do problema ao valor da multa é baixa. Este fato pode ser exemplificado pelos resultados da tabela 6: o nível de serviço com 7 caixas (75,2% dos clientes atendidos) é muito inferior ao nível de serviço com 8 caixas, 99,9%, tornando o custo de multas muito alto no cenário 7, mesmo para um valor de multa mais baixo.

	Número de caixas >					
	5	6	7	8	9	10
% de clientes atendidos em menos de 15 min	11,20	23,62	75,18	99,92	100	100
Custo total (R\$/mês)	438.129	333.455	113.394	20.629	22.446	24.940
				(ótimo)		

Tabela 6 – Resultados do modelo 1 com a multa valendo a metade do valor atual

	Número de caixas >					
	5	6	7	8	9	10
% de clientes atendidos em menos de 15 min	11,20	23,62	75,18	99,92	100	100
Custo total (R\$/mês)	1.715.073	1.288.927	401.204	22.659	22.446	24.940
					(ótimo)	

Tabela 7 – Resultados do modelo 1 com multa valendo o dobro do valor atual

A lei das filas determina que em dias de pagamento de funcionários públicos – dias de pico de atendimento – o tempo hábil deve ser inferior a 30 minutos. A tabela 8 apresenta os resultados da simulação do sistema representado pelo modelo 1 em dias de pico (os custos, nesse caso, são representados em R\$/dia). Considerou-se, neste caso, um aumento de 40% no movimento diário da agência.

	Número de caixas >							
	5	6	7	8	9	10	11	12
% de clientes atendidos em menos de 15 min	0,65	0,91	20,5	34,1	70,8	99,8	100	100
Custo total (R\$/dia)	46.562	46.532	34.962	26.229	12.102	1.259	1.219	1.330
Custo de multas (R\$/dia)	46.007	45.866	34.185	25.342	11.104	150	0	0

Tabela 8 – Resultados do modelo 1 considerando um dia de pico com 40% de aumento na demanda

A solução ótima em dias de pico exige 11 caixas de atendimento ou 3 caixas a mais em relação aos dias normais. Contudo, a solução com 10 caixas apresenta um custo médio de multas baixo e pode ser considerada pela agência como viável, caso existam mecanismos de atendimento alternativos no momentos em que a fila apresenta crescimento exagerado ao longo do dia.

5. Conclusões

Este artigo fez uma análise crítica da lei municipal 13.948/05, à luz da Pesquisa Operacional. As principais conclusões obtidas foram:

- a) O critério de tempo máximo de atendimento de 15 minutos é adequado como regra de critério de atendimento pois é facilmente entendido pelo usuário;
- b) O ponto de custo total (caixas e multas) ótimo, para os valores considerados na análise, é próximo do ponto de ótimo da agência, quando considerada a restrição de 15 minutos no tempo de atendimento;
- c) Quando considerados os custos usuais de operação bancária, o tempo de 15 minutos é compatível com o valor da multa adotada, forçando os bancos a operarem dentro do tempo e evitando-se assim, a situação em que pagar a multa é preferível ao custo de um novo caixa;
- d) A aplicação da lei aumenta o custo dos sistemas com filas para clientes preferenciais (idosos, gestantes e deficientes físicos) e especiais (categorizados assim por razões comerciais), o que pode provocar uma diminuição desta prática. Esta prática é injusta socialmente para o caso de clientes especiais, por desrespeitar a regra do primeiro a chegar é o primeiro a ser atendido. Contudo, para os clientes preferenciais – idosos, gestantes e deficientes físicos – o poder público deve agir no sentido de impedir que a lei se torne razão para o não atendimento preferencial desses clientes.

Como sugestão para trabalhos futuros, pode-se realizar uma cronoanálise mais apurada da movimentação em agências bancárias ao longo do dia. Espera-se que este estudo tenha contribuído no esclarecimento de parte das questões envolvidas na lei das filas, e oxalá os personagens envolvidos na gestão de sistemas com filas utilizem mais as ferramentas da Pesquisa Operacional no dimensionamento de sistemas que sejam satisfatórios para toda a população, independentemente do seu estado econômico ou social.

Referências

- BALBI, S.** *Bancos cobram para não atender clientes*. Folha de S. Paulo. 31 de outubro de 2005. Dinheiro, p.B8, 2005.
- CHWIF, L. & MEDINA, A.C.** *Modelagem e simulação de eventos discretos: teoria e aplicações*. 1. ed. São Paulo: Ed. dos Autores, 2006.
- CYPRIANO, M.** *Vencer o desafio das filas*. Folha de S. Paulo. 22 de setembro de 2005. Tendências/Debates, p.A3, 2005.
- KLEINROCK, L.** *Queueing Systems. Volume I: Theory*. 1. ed. New York: John Wiley & Sons, 1975.
- LARSON, R.** *Perspectives on queues: social justice and psychology of queueing*. Operations Research, v.35 (6), p.895-905, 1987.
- MORAES, V.F.** *Filas: responsabilidade e discurso*. Folha de S. Paulo. 11 de outubro de 2005. Tendências/Debates, p.A3, 2005.

SÃO PAULO. *Decreto nº 45.939, de 31 de maio de 2005.* Regulamenta a Lei nº 13.948, de 20 de janeiro de 2005. Secretaria do Governo Municipal. São Paulo, 31 de maio de 2005a.

SÃO PAULO. *Lei nº 13.948, de 20 de janeiro de 2005.* Dispõe sobre a obrigatoriedade de as agências bancárias e demais estabelecimentos de crédito de colocar, à disposição dos usuários pessoal suficiente no setor de caixas, para dar atendimento digno e profissional a seus clientes. Secretaria do Governo Municipal. São Paulo, 20 de janeiro de 2005b.