

# Controle de Congestionamento Veicular

R. Godoy, P. Castrucci

**Abstract**— A simple highway segment followed by a bottleneck is considered, subject to a wave or pulse of traffic demand, having restricted duration and a maximum which is higher than the capacity of the bottleneck. The proposed control of the flux is to impose a reduction of the maximum speed allowed in the segment, based on the measured arrival rate of vehicles (feedforward). The performance criterion is the weighted sum of the average traffic time with the average "stop and go" or queue time. The search for the optimal control rule is done through simulation with the software Arena; statistical distributions are "triangular".

**Index Terms**— vehicle traffic, highway, feedforward control, stochastic systems, simulation.

**Resumo**— Considera-se uma via expressa ("highway") de pista única com um gargalo, sujeita a uma onda ou pulso de demanda de tráfego, de duração restrita e de pico superior à capacidade do gargalo. O controle do fluxo baseia-se na imposição de redução na velocidade máxima permitida no trecho da via que precede o gargalo ("trecho pré-gargalo"), com base na medida da taxa de chegada de veículos (portanto "feedforward"). O critério de desempenho é a soma ponderada do tempo médio de percurso com o tempo médio de "parada" ou "stop-and-go" dos veículos. A pesquisa da regra de controle ótima é por simulação no software Arena; as distribuições estatísticas são triangulares.

**Palavras-chave** —trânsito de veículos; via expressa; controle em avanço; sistemas estocásticos; simulação.

## I. INTRODUÇÃO

O problema de trânsito considerado refere-se a uma via expressa ("highway") de pista única com *um gargalo*, causado por acidente ou obra em andamento. Supõe-se a via sujeita a *um surto de demanda de tráfego*, de forma conhecida, de duração restrita e de amplitude superior à capacidade do gargalo. O controle do fluxo baseia-se na imposição de *velocidade máxima intencionalmente reduzida*, no trecho da via que precede o gargalo; os motoristas são informados da sua limitação de velocidade por meio de painéis adequados, e coagidos a obedecer por monitoração eletrônica.

Este controle de velocidade no trecho que precede o gargalo tem por fim reduzir a taxa de chegada ao gargalo e portanto a fila correspondente e a duração de eventual processo de "stop and go"; o trecho que precede o gargalo age como uma espécie de "buffer" de veículos; além disso, para os

motoristas, deve ser interessante substituir o alto custo do "stop and go" desta fila por uma velocidade reduzida no trecho pré-gargalo.

O critério de eficiência adotado é a soma *ponderada* do tempo médio de percurso dos veículos do surto mais o tempo médio de permanência em fila de espera, em qualquer ponto da via; em fila de espera, há muito maior consumo de combustível, freios, cansaço, etc. Uma indicação da constante de ponderação pode ser obtida pelas taxas de cobranças vigentes nos táxis urbanos: o "minuto parado" custa o dobro do "km rodado". Supondo que o "km rodado" nas cidades se dê em velocidades da ordem de 30 km/h, o "km rodado" corresponde a 2 minutos; donde se extrai a indicação de que a constante de ponderação do critério deve ser da ordem de 4.

Vale recordar que um gargalo de tráfego funciona como um servidor de fila [1], com taxa de serviço, digamos,  $\mu$ ; sabe-se que se a taxa de chegada de clientes  $\lambda$  for maior ou igual a  $\mu$ , a esperança do comprimento da fila e a do tempo de atendimento tenderão ao infinito; uma eficiência razoável para qualquer serviço exige manter  $\lambda < \mu$ ; para casos simples, na prática, usam-se curvas-padrão de (comprimento ou espera na fila) versus  $(\lambda / \mu)$ . O problema proposto pode ser visto como o de ralentar adequadamente o tráfego antes do gargalo para que se minimize a probabilidade de atingir determinada taxa de chegada ao gargalo. Note que, em regime estacionário, com  $\lambda \geq \mu$ , o comprimento da fila no gargalo tende ao infinito e não existe hipótese de que alguma lei de controle das velocidades possa resolver o problema. Resta, e este é o objetivo da pesquisa, o caso não-estacionário de presença de uma onda de demanda de tráfego, de duração finita. Em outras palavras, só tem sentido considerar sobrecargas de trânsito que sejam transitórias.

O problema de congestionamento de vias expressas é tratado em várias publicações recentes, embora adotando uma variável de ação diferente; trata-se do controle do congestionamento em "highways" por meio de *semáforos instalados nas rampas de acesso* [3], [5]. Mede-se o congestionamento (densidade de tráfego, veic/km) imediatamente a jusante da rampa e se varia a relação verde-vermelho no semáforo da rampa; devido à proximidade medidor- semáforo considera-se não existir atraso puro na malha de controle; aplica-se então controle a realimentação, linear, discreto no tempo; os controladores são P+I.

Além das malhas locais [3], [2] consideram redes complexas de trânsito, com a mesma lei de controle local, incluindo a otimização global dos "set-points", por meio de programação matemática. (Papageorgiou e Kotsialos, 2000) informa que a otimização é importante para o caso de aparecimento imprevisto de gargalos. O método foi aplicado à perimetral de Amsterdã.

O problema tem conexão com o de telecomunicações por pacotes de dados [4], [5], [6].

O problema aqui considerado não permite solução semelhante, pois é impossível desprezar o atraso puro no processo; mais ainda, agir na variação da velocidade dos pelotões de veículos equivale a agir sobre o atraso puro do processo, no trecho pré-gargalo; em termos de controle linear seria ter como variável de controle o parâmetro T no expoente de uma função de transferência  $e^{-sT}$ . Em consequência, restringimo-nos à busca de uma lei ótima de controle do tipo "feedforward", que varia a velocidade máxima permitida no trecho, para cada lote de veículos que chega ao seu início.

Dada a natureza estocástica, a complexidade e a não-linearidade dos mecanismos de congestionamento e fila, o problema pertence ao difícil campo da *programação estocástica, em regime não estacionário*. Assim, para a solução, preferimos utilizar a simulação de eventos de sistemas discretos (Arena, Paragon / Rockwell Automation); por isso temos de nos restringir a casos específicos.

Considerou-se a chegada de um pulso ou "onda" de demanda, definida, discreta no tempo, típica de um início de dia na chegada às cidades. Em cada período de amostragem (1 minuto), a amplitude da "onda" é a média da distribuição estatística da taxa de chegadas de veículos (veic./min); a distribuição estatística dos intervalos de tempo entre chegadas sucessivas é exponencial. Para maior realismo, nos elementos da via adotaram-se distribuições de velocidade triangulares.

## II. O MODELO

### A. Sistema e Demanda

Considera-se um trecho de 3 km de via de pista única, do qual os 2990 m iniciais são destinados à velocidade controlada e às eventuais filas; no final do trecho (10 m) há uma obstrução parcial que leva a taxa máxima de serviço a 33 veículos/min.

Considera-se uma onda de demanda de tráfego que mantém média constante e igual a 66 veículos/min, durante 30 minutos. A distribuição dos intervalos de tempo entre veículos é suposta exponencial de média 0.015 minutos.

### B. Justificativa de uma regra de controle "feedforward"

Considere-se inicialmente a evolução do tráfego como um processo inteiramente determinístico que decorre de uma sucessão de impulsos de demanda de duração de 1 minuto, alguns abaixo e outros acima da taxa  $\mu g$ , escolhida como operação razoável do gargalo; o t-ésimo impulso é um lote de veículos, medido entre os instantes t-1 e t; suponha que ele se desloca na pista com uma velocidade recomendada, sem flutuações e sem espalhamento dos veículos uns em relação aos outros; por facilidade explicitam-se os atrasos recomendados dt, em vez das velocidades.

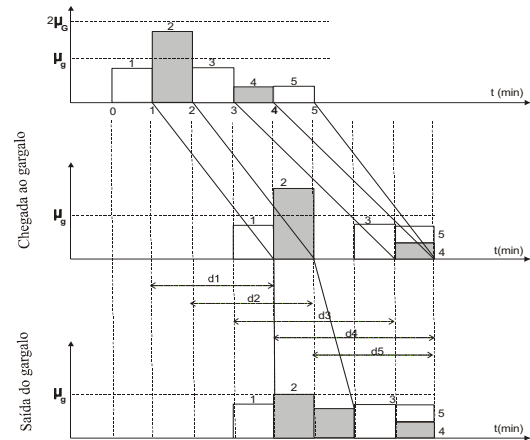


Fig. 1 - Processo de chegadas de dois lotes de veículos à pista

Observe a figura 1, especialmente os lotes 2 e 3: desejando reduzir as filas de espera no gargalo, é evidente que o atraso dt deve ser igual ao atraso do lote anterior dt-1, mais um certo número de períodos suficiente para que o lote em t se escoe pelo gargalo, a uma taxa de  $\mu g$  veic/min.

Seja  $MI(\text{lotet}/\mu g)$  o maior inteiro contido em  $(\text{lotet}/\mu g)$ ; em primeira aproximação pode-se adotar como regra de controle:

- I. Se  $\text{lotet} > \mu g$   
 $dt = dt-1 + MI(\text{lotet}/\mu g)$
- II. Se  $\text{lotet} \leq \mu g$   
 $\text{lotet}-1 + \text{lotet} \geq \mu g$   
 $dt = dt-1$
- III. Se  $\text{lotet} \leq \mu g$   
 $\text{lotet}-1 + \text{lotet} < \mu g$   
 $dt = dt-1-1$

Na realidade o processo de passagem pelo trecho anterior ao gargalo introduz um espalhamento aleatório dos veículos de cada lote; além disso, quando um lote  $> \mu g$  é limitado pelo gargalo, ele estende sua ocupação do mesmo por mais de um período, o processo tem naturalmente aumentadas as suas incertezas. Por estas razões, adotamos a regra I', II' em vez de I, II, III.

- I'. Se  $\text{lotet} > \mu g$   
 $dt = dt-1 + \varepsilon * MI(\text{lotet}/\mu g)$
- onde  $\varepsilon$  é um fator a ajustar empiricamente, otimizando o desempenho por simulações,

seguida da regra II' em vez de II e III,

- II'. Se  $\text{lotet} \leq \mu g$   
 $dt = dt-1-1$

Salienta-se que  $\varepsilon$  deve ser ajustado como um parâmetro do controlador PID. Uma vez ajustado, para cada caso específico de via, proporciona o controle "on-line" do fluxo de tráfego, uma vez que o gerenciamento de tráfego é feito pela contagem

dos lotes de veículos que chegam ao início do trecho de pista, a cada minuto.

C. Critério de desempenho

[7] adotaram os indicadores VMT e VKT definidos a seguir, e a interessante visão de que uma via é uma "máquina" que "consome" VMT veículos x minuto e que "produz" VKT veículos x km. No presente caso, VKT é constante, dados o surto de veículos e a extensão da pista; pode ser omitida no critério de desempenho. No presente caso, o VMT precisa ser modificado para considerar os custos diferentes dos veículos rodando e parados.

Tudo considerado, decidiu-se trabalhar com o critério de desempenho:

$$J = vmm + \rho \quad \times \quad vmpm \quad (4)$$

onde

vmm = valor médio do tempo de travessia = VMT/NC

sendo

$$VMT = \sum_{i=1}^{NC} TT_i \quad (5)$$

NC = número de veículos que percorreram a pista durante a simulação;

TT<sub>i</sub> = tempo de travessia do trecho completo, pelo i-ésimo veículo

$$vmpm = wt(q1) + wt(q2) \quad (6)$$

onde

wt(q1) = tempo de parada médio na fila q1, situada na entrada do trecho de 3 km

wt(q2) = tempo de parada médio na fila q2, situada na entrada do gargalo de 10 m

ρ = constante de ponderação, sócio-econômica = 5

III. SIMULAÇÃO NO ARENA

De acordo com a prática e com várias teorias conhecidas, em qualquer elemento de via a velocidade média efetiva v (km/h) e o fluxo médio efetivo f (veic/h) são funções não-lineares da velocidade máxima permitida e da densidade do tráfego k (veículos/km) [10]. [1] justifica-as plenamente a partir de uma descrição do elemento e do trecho que o precede, como servidor e fila, respectivamente. É esta descrição que servirá de base para a simulação.

A via expressa em estudo considera-se decomposta em 2 trechos sucessivos: 2990 m de "buffer" de veículos em velocidade controlável e 10 m de gargalo.

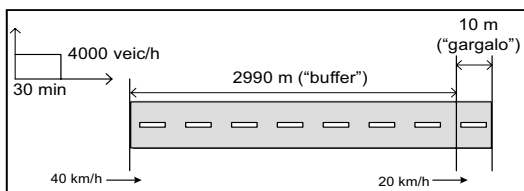


Fig. 2 - Via expressa de 3 km

A figura 3 refere-se à simulação do trecho maior, compreendendo os seguintes módulos Arena:

- *Seize pista2990*: representa a ocupação pela entidade veículo do recurso trecho de pista de 2990m;
- *Delay pista2990*: introduz um atraso de distribuição triangular de valores mínimo, médio e máximo dados: TRIA(min, med, max) = TRIA(0.8\*(4.5+(299/300)\*atraso), (4.5+(299/300)\*atraso), 1.2\*(4.5+(299/300)\*atraso)) (7), onde atraso = valor calculado de dt por (2) (3) em minutos; 4.5 minutos = tempo para percorrer 2990 m a 40 km/h;
- *Release pista2990*: proporciona a desocupação do recurso trecho. Nota-se que, associado ao bloco Seize, tem-se a fila q1 de veículos parados, que não cabem no trecho.

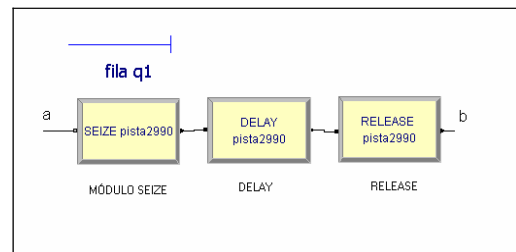


Fig 3 - Diagrama Arena para trecho de 2990 metros

A Figura 4 refere-se à chegada e contagem dos lotes de demanda lotet.

Descrição dos módulos Arena da Figura 4:

- *Create*, chegada de veículos: injeta no modelo as entidades que representam os veículos, cujo intervalo entre chegadas é dado por um schedule de 30 minutos de duração, a uma taxa de 66 veículos/min, equivalente a um total de 2000 veículos;
- *Assign*, contador geral de carros: implementa a variável contadora "entrada";
- *Decide*,  $x \geq 2000$ ?: verifica se a variável contadora "entrada"  $> 2000$ ; caso seja, faz-se dt igual a zero e terminam-se os cálculos de dt;
- *Assign*, contador de carros por minuto na entrada da pista: implementa a variável contadora de veículos por minuto, xi;
- *Assign*, totalizador de carros na entrada da pista: implementa a variável contadora "entradas";
- *Assign, atraso em minutos*: define o atributo de entidade "atraso", numericamente igual à variável d<sub>i</sub>;
- *Assign, cronometro 01*: faz a tomada do instante de entrada de cada entidade no sistema.

A figura 5 trata do diagrama Arena para o trecho gargalo de 10 m e da coleta de estatísticas de simulação.

Descrição dos módulos Arena da Figura 5:

- *Server*, gargalo: refere-se ao "processamento" de veículos pelo recurso gargalo\_R, sob um tempo de processo igual a TRIA(0.8\*0.03, 0.03, 1.2\*0.03), tendo associada a fila q2 de veículos; 0.03 minutos é o tempo para percorrer 10 m a 20 km/h;

controlado. Quanto menor o J, melhor o desempenho. Ver Figura 6 para o desempenho J no total das simulações.

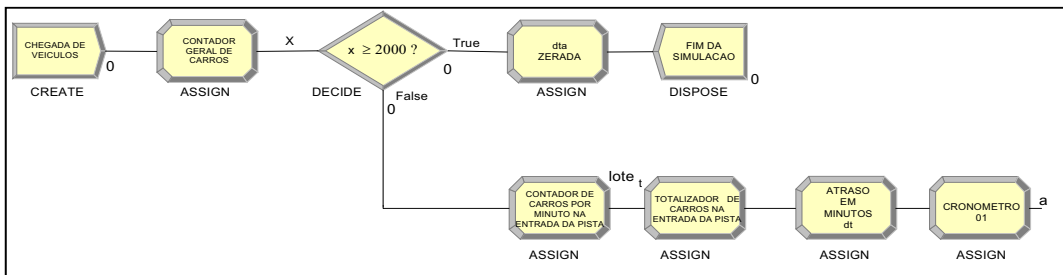


Fig 4 - Diagrama da chegada e contagem dos lotes de demanda lote<sub>i</sub>, omitido o relógio de amostragem

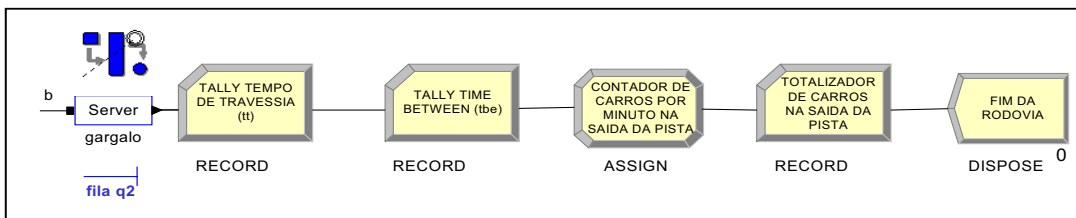


Fig 5 - Diagrama do trecho gargalo e da coleta de estatísticas de simulação

- *Record tally*, tempo de travessia: grava no totalizador "tt" o tempo de travessia para os 3 km de pista;
- *Record tally*, time between: grava no totalizador "tbe" o tempo entre chegadas de veículos na saída;
- *Assign*, contador de carros por minuto na saída da pista: implementa a variável contadora de carros por minuto, xo;
- *Record*, totalizador de carros na saída da pista: implementa o contador "saída";
- *Dispose*, fim da rodovia: representa a saída de veículos da pista, correspondendo à eliminação de entidades no modelo de simulação.

Um outro diagrama, aqui omitido, basicamente recebe como entrada a contagem do número de veículos por minuto (amostragem) na entrada da rodovia, xi, e calcula como saída o atraso dt a ser aplicado a cada entidade veículo, de acordo com as regras de controle (I') e (II').

IV. RESULTADOS

As tabelas a seguir mostram o desempenho da rodovia em algumas condições da simulação. As simulações foram replicadas 20 vezes; as incertezas da tabela são meias larguras dos intervalos de confiança, calculados pelo Arena, sendo o tally " tbe " = intervalo de tempo médio entre a passagem de dois veículos sucessivos.

Note-se que vmm é igual ao tally " tt " (tempo de travessia para 3 km), e vmpm é igual a wt(q1)+wt(q2). O critério básico é J=(vmm+5\*vmpm) para o sistema controlado e não-

Tabela 1 - Resultados de simulações para 20 replicações

	ε=0.05	ε=0.10	ε=0.15
vmm=tt	(19.33±0.20)	(19.22±0.21)	(19.30±0.15)
tbe	(0.03±0.00)	(0.03±0.00)	(0.03±0.00)
wt(q1)	(2.05±0.16)	(3.81±0.20)	(5.78±0.22)
wt(q2)	(11.61±0.11)	(8.65±0.16)	(5.64±0.26)
vmpm	(13.66±0.19)	(12.46±0.26)	(11.42±0.34)
q1	(63.82±5.05)	(118.33±6.15)	(179.78±6.79)
q2	(360.96±3.27)	(268.69±4.98)	(175.17±8.27)
J	(87.63±0.97)	(81.52±1.32)	(76.40±1.71)

	ε=0.18	ε=0.20	ε=0.22
vmm=tt	(19.27±0.21)	(19.57±0.24)	(20.36±0.20)
tbe	(0.03±0.00)	(0.03±0.00)	(0.03±0.00)
wt(q1)	(6.87±0.27)	(7.59±0.23)	(8.57±0.16)
wt(q2)	(3.87±0.25)	(3.03±0.13)	(2.33±0.11)
vmpm	(10.74±0.37)	(10.62±0.26)	(10.90±0.19)
q1	(204.47±6.53)	(215.13±5.39)	(230.34±3.31)
q2	(115.60±8.47)	(86.02±4.12)	(62.61±3.41)
J	(72.97±1.86)	(72.67±1.32)	(74.86±0.97)

	ε=0.25	ε=0.30	sem controle
vmm=tt	(21.41±0.24)	(23.92±0.32)	(19.23±0.18)
tbe	(0.03±0.00)	(0.04±0.00)	(0.03±0.00)
wt(q1)	(9.62±0.20)	(11.65±0.26)	(0.18±0.08)
wt(q2)	(1.70±0.08)	(1.00±0.08)	(14.52±0.12)
vmpm	(11.32±0.22)	(12.65±0.27)	(14.70±0.14)
q1	(242.68±4.31)	(262.86±3.63)	(5.61±2.33)
q2	(43.07±2.17)	(22.71±1.90)	(451.71±3.59)
J	(78.01±1.13)	(87.17±1.39)	(92.73±0.72)

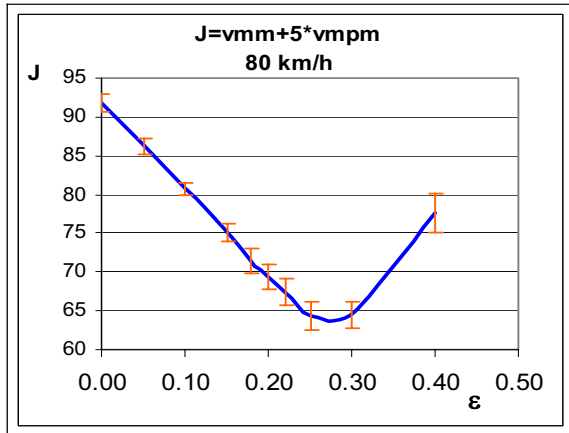


Fig. 6 - Desempenho J em função de ε; ε = 0 é caso sem controle

V. CONCLUSÕES

O método de controle "feedforward" apresentado é confirmado pelas simulações, e o ε ótimo é nítido. Quando se aumenta o surto de demanda, o mínimo J pode desaparecer; isto ocorre quando a capacidade do pré-gargalo como "buffer" fica insuficiente. Note que com 3 km o máximo de veículos retidos é 300. Esta é a principal limitação prática.

Questões a pesquisar: a) efeitos de diferentes parâmetros da onda de demanda, sobre o ε ótimo; b) avaliar custo do investimento versus benefício sócio-econômico.

VI. ANEXO MODELO DE HEIDEMANN

Considere-se uma rodovia com tráfego fluindo ininterruptamente numa direção. Não há intersecções ou outros obstáculos, de modo que os problemas que podem surgir são advindos do próprio fluxo de tráfego. Assumem-se condições estacionárias, isto é, fluxo de tráfego em equilíbrio estatístico. A notação a seguir é usada:

- k é a densidade de tráfego (veic/km)
- v é a velocidade individual ou a velocidade média (km/h)
- kjam é a densidade máxima ou densidade de congestionamento (isto é, a mais baixa densidade sob a qual o fluxo de tráfego chega a uma parada)
- vf é a velocidade individual desejada ou a velocidade média desejada
- q é a taxa de fluxo (veic/h).

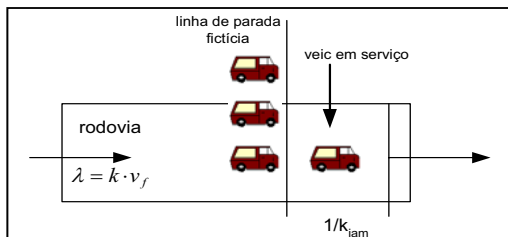


Fig 7 - Modelo de rodovia particionada em trechos de comprimento 1/kjam

Imaginemos a situação fictícia e ideal em que cada veículo é conduzido sob a velocidade desejada e numa densidade qualquer de tráfego. Particiona-se a rodovia em trechos de comprimento 1/kjam. Em virtude da definição de kjam este é o comprimento mínimo de rodovia de que cada veículo necessita. O tempo médio para atravessar este trecho de rodovia é 1/(kjam x vf), se os veículos viajam sob a velocidade desejada. Então, a correspondente taxa de fluxo é μ = kjam x vf. A seguir, imaginamos um trecho arbitrário de rodovia de comprimento 1/kjam como sendo um servidor operando sob taxa de serviço μ, e sabemos que esta é uma taxa desejada fictícia. A taxa de serviço deste servidor é λ = k x vf, desde que a densidade k prevaleça e que a velocidade do veículo seja a desejada. Além disso, assume-se que as chegadas constituem um processo de Poisson, isto é, os intervalos de tempo entre chegadas são exponencialmente distribuídos. Devido ao fenômeno de fila usual os veículos que chegam podem ter de formar uma fila em frente ao servidor porque este pode estar ocupado. Imaginamos uma linha de parada defronte ao servidor onde os veículos aguardam por serviço, e a linha é suposta suficientemente longa para acomodar todos os veículos em fila.

Considerando o trecho 1/kjam como um servidor da teoria das filas, o autor deduz as relações tradicionais entre velocidade média v (km/h), congestionamento k veic/km e fluxo q veic/h, bem como verifica sua validade por medidas de campo.

$$v = \frac{2 \times v_f \times (1 - \eta)}{2 + \eta \times (\beta^2 - 1)} = v(\eta)$$

onde:

$$\eta = \frac{k}{k_{jam}} = \text{"grau de saturação" da rodovia}$$

$\beta = \sigma \times v_f \times k_{jam}$  = coeficiente de variação do tempo de serviço, que é o tempo (com desvio-padrão σ) para percorrer a distância 1/kjam sob a velocidade desejada.

$$q = k \times v(\eta) = \eta \times k_{jam} \times v(\eta)$$

VII. REFERENCES

[1] [1] Heidemann, D. (1996). A queuing theory approach to speed-flow-density relationships, Transportation and Traffic Theory, Pergamon.  
 [2] [2] Benmohamed, L. and Meerkov, S.M. (1994). Feedback control of highway congestion by a fair on-ramp metering, Proc 33rd Conf on Decision and Control.  
 [3] [3] Papageorgiou and Kotsialos, A. (2000). Freeway ramp metering: an overview, IEEE Intelligent Transport Systems Conf Proc.  
 [4] [4] Benmohamed, L. and Meerkov, S.M. (1993). Feedback control of congestion in packet switching networks, IEEE Trans on Networking.  
 [5] Benmohamed, L. and Meerkov, S.M. (1994). Feedback control of congestion in packet switching networks, Proc American Control Conference.  
 [6] Gerla, M. and Kleinrock, L. (1980). Flow control: a comparative survey, IEEE Trans on Communicatios.

- [7] Chao, C. and Varaiya, P. (2001). Causes and Cures of Highway Congestion, IEEE Control Systems Magazine.
- [8] Kelton, W. D., Sadowski, R.P. and Sadowski, D.A. (1998). Simulation with Arena, McGraw-Hill
- [9] Arena User's Guide, Rockwell Software Inc., Sewickley, PA, USA.
- [10] Papacostas, C.S. (1987). Fundamentals of Transportation Engineering, Prentice Hall

### VIII. BIOGRAFIAS



Plínio Castrucci foi professor titular da área de Controle Automático, no Departamento de Engenharia Elétrica da Escola Politécnica da Universidade de São Paulo (1967-1987). Graduiu-se em Engenharia Mecânica e Elétrica, obteve os títulos de D.I.C. no Imperial College of Science and Technology (1958) e de livre-docente e professor titular na EPUSP. Foi diretor da Amplimag Controles

Automáticos S.A. e diretor superintendente da Engesa Elétrica S.A. Tem sido consultor de empresas. É senior member IEEE e autor dos livros: Controle Automático – teoria e projeto, Controle Linear – método básico, Sistemas Não Lineares, Controle Digital.



Roberto Godoy obteve o título de engenheiro eletricista pela Escola Politécnica da Universidade de São Paulo em 1991, e de mestre (1999) e doutor (2005) em Engenharia Elétrica também pela EPUSP.

Trabalhou na área de gerenciamento de obras bancárias e, desde 1997, atua como engenheiro pesquisador no Convênio Rockwell-EPUSP. Entre seus interesses estão o modelamento e simulação de sistemas a eventos discretos e informática.