

# Um Esquema Bio-inspirado para Estimação de Campo com Redes de Sensores sem Fio

Daniel de O. Cunha, Otto Carlos M. B. Duarte\*

<sup>1</sup>Grupo de Teleinformática e Automação  
PEE/COPPE - DEL/POLI

Universidade Federal do Rio de Janeiro

C.P. 68504 - CEP 21945-970

Rio de Janeiro - RJ - Brasil

<http://www.gta.ufrj.br>

{doc,otto}@gta.ufrj.br

**Abstract.** *This paper proposes and analyzes a bio-inspired scheme to field estimation using wireless sensor networks. The proposed scheme exploits the temporal pattern of the sensed process to reduce the number of samples sent back by a sensor node to the sink and decrease the energy consumption in data transmission. The proposed scheme is orthogonal to the techniques that reduce the spatial density of collected samples deactivating nodes with similar measurements. Thus, the proposed scheme can be used along with these techniques.*

**Resumo.** *Este artigo propõe e analisa um esquema bio-inspirado para estimação de campo utilizando redes de sensores sem fio. O esquema proposto explora características temporais do processo sensoriado para reduzir o número de amostras enviadas, ao longo do tempo, por nós sensores ativos ao sorvedouro e, conseqüentemente, diminuir o consumo de energia com a transmissão de dados. O esquema proposto é ortogonal às abordagens convencionais que buscam reduzir a densidade espacial das amostras coletadas desativando nós com medidas redundantes, podendo ser utilizado em conjunto com essas técnicas.*

## 1. Introdução

Os recentes avanços na microeletrônica e na comunicação sem fio permitiram o desenvolvimento de nós sensores de baixo custo capazes de comunicarem-se a curtas distâncias. Estes pequenos sensores, dotados de unidades de sensoriamento, processamento e comunicação, podem ser organizados numa complexa estrutura cooperativa, dando origem a uma rede de sensores sem fio [Akyildiz et al., 2002]. O baixo custo individual destes nós torna viável a construção de redes de sensores compostas por centenas ou milhares de nós. Isto possibilita uma maior proximidade desta rede com o ambiente monitorado, possibilitando a obtenção de dados mais detalhados que dificilmente poderiam ser obtidos por técnicas de sensoriamento convencionais, como satélites e radares.

Uma das aplicações mais importantes das redes de sensores sem fio é a estimação de campo. Nestas aplicações são espalhados diversos nós sensores numa determinada área para fazer o sensoriamento remoto de processos variáveis espaço-temporalmente nesta região. A qualidade da estimativa a cerca do comportamento do campo sensoriado está diretamente ligada à frequência espaço-temporal com que são obtidas as

---

\*Este trabalho foi realizado com recursos do CNPq, CAPES, FAPERJ, RNP, FINEP e FUNTTEL.

amostras. A Frequência espacial está diretamente ligada ao número de nós sensores e à forma com que estes sensores são colocados na região, enquanto que a frequência temporal depende da taxa com que os nós ativos coletam os dados. Quanto maior esta frequência, maior será a precisão na reconstrução do comportamento do processo monitorado [Kumar et al., 2004]. No entanto, um maior número de amostras implica uma maior sobrecarga de comunicação para a transmissão dos dados na rede e, conseqüentemente, um maior consumo de energia por parte dos nós sensores.

A solução mais difundida para a redução desta sobrecarga na rede é identificar locais com leituras semelhantes e diminuir a densidade espacial da coleta de amostras [Willett et al., 2004, Nowak et al., 2004, Rahimi et al., 2004] nessas regiões. Para isto, nós sensores medindo valores muito próximos uns dos outros são considerados redundantes e desativados. Desta forma, evita-se a transmissão de dados que acrescentem pouca ou nenhuma informação ao processo de estimação do campo. Esta redução pode aumentar significativamente o tempo de vida dos nós sensores, pois a transmissão de dados é, em geral, relativamente mais custosa do que o processamento realizado para detectar os pontos de redundância na rede [Pottie and Kaiser, 2000].

Há, no entanto, uma outra dimensão a ser considerada na coleta de dados: a densidade temporal das amostras. Enquanto muitos trabalhos buscam reduzir o número de amostras coletadas com a desativação de nós sensores redundantes, como discutido anteriormente, pouco esforço tem sido feito no sentido de determinar quando os nós que permanecem coletando dados, ou ativos, devem enviar suas leituras ao nó sorvedouro. Uma possível redução na frequência temporal de coleta de amostras pode funcionar de forma ortogonal à redução espacial e, de forma similar a esta última, deve ser controlada com base no erro aceitável para a estimativa do campo. Alguns trabalhos tratam da codificação de séries temporais através de diferentes técnicas, como a PCA (*Piecewise Constant Approximation*) [Lazaridis and Mehrotra, 2003] e wavelets [Chen et al., 2004].

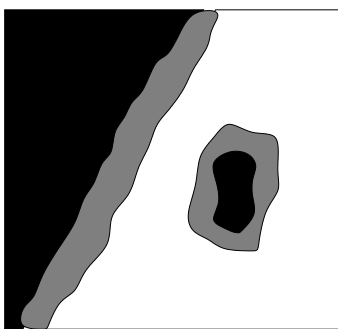
Uma característica importante que pode ser aproveitada é o comportamento regular que muitos processos físicos apresentam. A temperatura, por exemplo, exibe em geral um comportamento similar ao longo de dias consecutivos. Uma possibilidade para explorar estas características é fazer com que os nós sensores identifiquem padrões no processo físico e apenas reportem ao sorvedouro situações diferentes do comum. Esta percepção do ambiente é similar à forma com que os seres vivos respondem às ocorrências ao seu redor. Tanto pessoas como animais estão constantemente sujeitos a uma série de estímulos. No entanto, é impossível tratar conscientemente de todos estes estímulos. Os organismos desenvolvem então a noção de periferia e centro da atenção [Weiser and Brown, 1997]. A periferia é composta pelos estímulos levados em consideração de forma subconsciente, enquanto que o centro da atenção é o elemento conscientemente tratado num dado instante. Em geral um elemento sai da periferia para o centro quando ocorre um evento fora do comum, destoante dos outros estímulos periféricos.

Este trabalho propõe e analisa um esquema bio-inspirado para explorar as características específicas dos processos monitorados e reduzir o número de amostras enviadas ao nó sorvedouro. O esquema proposto funciona de forma distribuída e baseia-se na construção por parte de cada um dos nós sensores de suas próprias periferias de atenção. Desta forma os nós somente reportam ao sorvedouro as amostras que se destaquem em relação ao normalmente observado pelo nó. Este artigo está organizado da seguinte forma. A seção 2 apresenta alguns trabalhos relacionados à estimativa de campo. A Seção 3 demonstra mais claramente como algumas características dos processos físicos podem ser exploradas e detalha o esquema proposto. Na Seção 4 são apresentados os resultados das simulações do esquema proposto. Por fim, a Seção 5 conclui este trabalho.

## 2. Trabalhos Relacionados

Muitos trabalhos sobre estimação de campos preocupam-se com questões relacionadas à frequência espacial com que as amostras são coletadas. A maior parte dos trabalhos considera campos com banda limitada, ou seja, cuja variação no comportamento entre dois pontos é limitada a uma determinada frequência. No entanto, a estimação de campos sem limite de banda é viável [Kumar et al., 2004]. Kumar *et al.* sustentam que o erro causado pela sub-amostragem espacial não precisa ser eliminado, apenas reduzido ao nível do erro de quantização do sensor. Além disso, eles demonstram que há um compromisso entre a densidade de sensores e sua precisão para a obtenção de um determinado erro na reconstrução do campo.

A maior parte dos trabalhos que buscam reduzir a frequência espacial de coleta de dados assume a existência de um campo a ser sensoriado similar ao mostrado na Figura 1. Nesta figura, a região preta e as regiões brancas representam áreas onde o processo sensoriado varia suavemente em função da posição do sensor. As regiões cinza, por outro lado, representam áreas onde o processo apresenta uma variação mais abrupta em função da posição do nó sensor e são chamadas de fronteiras.



**Figura 1: Exemplo de campo com variações suaves e poucas áreas de fronteira.**

O esquema conhecido como *Backcasting* [Willett et al., 2004] propõe uma abordagem adaptativa em dois passos para reduzir o número de sensores ativos em função das correlações existentes no ambiente. Numa primeira etapa, um subconjunto dos nós sensores faz uma estimativa inicial do campo. Com base nesta estimativa, o centro de controle ativa mais nós sensores em regiões que apresentem maior variação nas medidas. Desta forma, regiões com um comportamento mais suave serão monitoradas com uma densidade menor de nós, enquanto que regiões de fronteira são monitoradas com uma maior densidade de nós sensores. Esta abordagem privilegia a coleta de dados nas regiões com mudanças abruptas no campo, que tendem a ser de maior interesse [Nowak et al., 2004]. Apesar de conseguir reduzir o número de amostras coletadas nas regiões do campo com variação mais suave, este esquema não é capaz de reduzir o número de amostras nas regiões de fronteira, onde a rede tende a operar com a maior densidade possível de nós sensores ativos.

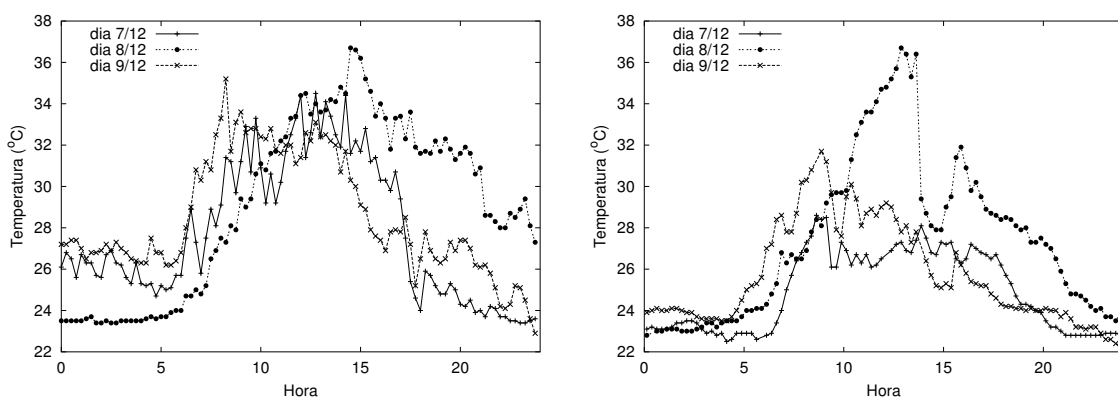
A proposta do algoritmo *Fidelity Driven Sampling* - FDS [Rahimi et al., 2004, Batalin et al., 2004] é explorar uma possível mobilidade para fazer uma estimativa inicial do ambiente e dividi-lo em regiões de acordo com a variação das medidas coletadas. Com base nesta estimativa, são adicionados pontos de amostragem às regiões com maior variação até que o erro na estimativa chegue no patamar desejado ou que a cota máxima de amostras que podem ser coletadas seja atingida. Este esquema, no entanto, sofre da mesma limitação da técnica *Backcasting* por só preocupar-se com a densidade espacial das medidas. Além disso, é necessária a utilização de uma infra-estrutura de cabos para prover a mobilidade necessária aos sensores.

O algoritmo RACE [Chen et al., 2004] codifica de forma adaptativa as amostras coletas através do uso de uma transformada wavelet. Desta forma, consegue-se reduzir o total de dados transmitidos ao sorvedouro. É possível neste tipo de abordagem a codificação das amostras com outras técnicas, como a PCA (*Piecewise Constant Approximation*) [Lazaridis and Mehrotra, 2003]. Estas técnicas, no entanto, são utilizadas após a decisão sobre quais amostras serão transmitidas e não reduzem o número de amostras enviadas. Elas transmitem menos dados por realizar uma codificação mais eficiente das amostras.

O esquema proposto neste trabalho visa complementar estas técnicas de redução na densidade espacial de coleta de dados e codificação mais eficiente das amostras. O objetivo desta proposta é reduzir a densidade temporal das amostras, explorando características intrínsecas do processo monitorado. Por ser uma técnica que atua em cada nó sensor individualmente, esta técnica possibilita até mesmo a redução do número de amostras geradas pelos nós nas regiões de fronteira, economizando recursos onde técnicas como o *Backcasting* e o FDS não conseguem economizar. Além disso, a técnica proposta possui a vantagem de ser distribuída, pois cada nó sensor define sua própria idéia de periferia, como será discutido na Seção 3, e dispensa a utilização de qualquer infra-estrutura para operação.

### 3. Esquema Proposto

O esquema proposto baseia-se na construção de uma idéia de periferia de atenção por parte de cada nó sensor. Para que isto seja possível, é necessário que o processo monitorado apresente um comportamento ao longo do tempo com alguma característica que possa ser considerada comum, ou esperada. É importante destacar que muitos processos físicos apresentam um comportamento regular, às vezes com ciclos bem definidos. A Figura 2 mostra a temperatura coletada ao longo de três dias consecutivos por duas estações meteorológicas distintas na cidade do Rio de Janeiro [GEORIO, 2003].



(a) Temperatura em São Cristóvão.

(b) Temperatura em Guaratiba

**Figura 2: Temperaturas em dois lugares distintos da cidade do Rio de Janeiro.**

Analisando esta figura, fica claro um comportamento padrão onde a temperatura é inicialmente mais baixa pela manhã e começa a subir próximo ao meio-dia. A temperatura declina no meio da tarde e chega a níveis relativamente baixos à noite. Considerando a idéia de periferia discutida anteriormente, uma medida que apresente um comportamento similar ao comportamento médio destas curvas poderia permanecer na periferia, enquanto

que algum comportamento diferindo além de um certo limiar deste comportamento esperado seria levado ao centro da atenção por parte do nó sensor. No exemplo citado, pode-se perceber que as curvas da Figura 2(a) apresentam um comportamento bem parecido. As curvas da Figura 2(b), no entanto, mostram um comportamento destoante no dia 8/12, especialmente a partir das 10 horas. O esquema proposto visa exatamente economizar energia não mandando amostras no momento em que o comportamento da medida for o esperado e enviar amostras com maior frequência em situações como a mostrada na Figura 2(b). Desta forma, o sorvedouro assume que o comportamento da medida segue o esperado quando não receber nenhuma mensagem de atualização de amostra e ao receber os dados referentes ao comportamento incomum do dia 8/12, é capaz de estimar o campo sensoriado com pouco erro levando em consideração os comportamentos incomuns, como o mostrado para o dia 8/12 na Figura 2(b). Estas amostras enviadas devido a uma diferença grande em relação ao comportamento esperado são chamadas de amostras de refinamento.

Este processo de construção de um comportamento esperado e definição sobre amostras destoantes pode ser feito de diferentes formas. Uma etapa inicial é identificar a periodicidade com que o comportamento regular ocorre. Uma opção é a coleta de dados com taxa máxima de amostragem por um determinado período maior do que a regularidade esperada e o cálculo da autocorrelação deste vetor com as medidas obtidas. Com base neste cálculo de autocorrelação pode-se obter a periodicidade do processo. Uma vez conhecida esta periodicidade o nó deve começar a identificar o comportamento periódico esperado. O nó pode obter este comportamento periódico esperado da seguinte forma:

$$CP_i = P_{i-1} \times \alpha + CP_{i-1} \times (1 - \alpha), \quad (1)$$

onde  $CP_j$  é um vetor contendo o comportamento esperado ao longo do período  $j$ ,  $P_{i-1}$  é o vetor com as medidas obtidas ao longo do período  $(i - 1)$  e  $\alpha$  é um fator multiplicativo que determina o peso da última amostra no cálculo do comportamento esperado.

Desta forma, quanto maior  $\alpha$  maior o peso das amostras recentes e menos valor terá o histórico do comportamento do processo. Por outro lado, um  $\alpha$  pequeno dá maior valor ao histórico do processo e valoriza pouco variações súbitas. Com base na Equação 1, pode-se verificar que o peso ( $W$ ) das medidas obtidas a  $i$  períodos no comportamento esperado é:

$$W = \alpha \times (1 - \alpha)^{i-1} \mid i \geq 1. \quad (2)$$

No caso da temperatura, medida abordada mais detalhadamente neste trabalho, há claramente uma periodicidade diária. Outras periodicidades podem ser identificadas, como por exemplo anual, mas as análises realizadas neste trabalho são focadas na periodicidade diária. Desta forma, os nós sensores devem identificar um comportamento diário esperado, que vai sendo atualizado a cada dia. A decisão sobre que amostras enviar é feita com base neste comportamento. Há no entanto uma questão prática ligada a como o sorvedouro será capaz de reconstruir a informação enviada pelo nó sensor. Para que isto seja possível, o nó deve enviar periodicamente um comportamento esperado ao sorvedouro. Este comportamento esperado deve servir de base para a reconstrução da informação fornecida pelo sensor. Até que um próximo comportamento esperado seja enviado pelo sensor, o sorvedouro assume que a informação coletada pelo nó segue o último comportamento esperado, a exceção das amostras de refinamento enviadas pelo nó sensor que devem ser colocadas no dia e horário especificados pelo sensor. O nó sensor deve fazer a análise dos dados a serem enviados com base no último vetor de comportamento esperado enviado ao sorvedouro, de modo a garantir que o sorvedouro esteja reconstruindo a informação como o nó sensor espera. Assim, o nó sensor pode averiguar se a medida

obtida para um dado instante difere muito ou pouco do valor enviado no último vetor com comportamento esperado para aquele horário específico. Caso a diferença, ou erro, em relação ao valor esperado for menor que um certo limiar configurável na rede, o nó sensor não envia a amostra ao sorvedouro, se esta diferença for maior que o erro admitido na reconstrução, o nó sensor envia a amostra juntamente com a informação temporal para inserção desta amostra na estimação do campo no sorvedouro.

Duas questões devem ser destacadas quanto a esse procedimento. Primeiramente, deve-se destacar que embora as medidas devam ser comparadas com o comportamento esperado para um determinado horário, este mecanismo não requer nenhum tipo de sincronização entre os diferentes nós sensores, pois cada sensor constrói sua própria periferia e analisa as amostras numa base de tempo local. O esquema funciona de uma forma distribuída. Além disso, se a cada dia o nó enviar o comportamento esperado para o dia seguinte, o sistema não obterá nenhum ganho. Desta forma, o sensor deve enviar as atualizações de comportamentos esperados a intervalos que possibilitem redução no número de amostras enviadas. O procedimento realizado diariamente pode ser visto na Figura 3, onde  $CD_j$  é um vetor contendo o comportamento esperado ao longo do dia  $j$ ,  $D_i$  é o vetor com as medidas obtidas ao longo do dia  $i$ , *última\_atualização* é o último vetor de atualização do comportamento esperado enviado e  $X(k)$  é o  $k$ -ésimo elemento do vetor  $X$ .

$$CD_i = \alpha D_i + (1-\alpha) CD_{i-1}$$

Se momento de atualização então

*última\_atualização* =  $CD_i$

Enviar *última\_atualização*

Senão

Para todas as  $k$  amostras em  $D_i$  faça

Se  $|D_i(k) - \text{última\_atualização}(k)| > \text{última\_atualização}(k) * \text{erro\_configurado}$

Enviar  $D_i(k)$

**Figura 3: Procedimento realizado diariamente por cada nó sensor.**

O cálculo do comportamento esperado no algoritmo mostrado na Figura 3 é ligeiramente modificado em relação ao mostrado na Equação 1 para enviar, no dia em que é feita a atualização periódica, o comportamento esperado para este dia (*última\_atualização* =  $CD_i$ ) no lugar das medidas obtidas neste dia ( $D_i$ ). Esta é apenas uma simplificação, mas o algoritmo poderia mandar tanto a atualização quanto as amostras com diferença maior que a permitida no dia em que é feita a atualização. À medida que o valor de  $\alpha$  aumenta, e a variação do processo de um dia para o outro diminui, estas duas abordagens tendem a se igualar. A utilização do algoritmo mostrado na Figura 3, que baseia-se na Equação 1 para a obtenção do comportamento esperado, torna o esquema de baixa complexidade computacional. Esta baixa complexidade viabiliza a utilização do esquema proposto em nós sensores muito simples e com baixo poder computacional. A próxima seção detalha as simulações realizadas para analisar o esquema proposto.

## 4. Simulações

Com o intuito de verificar a viabilidade do esquema proposto, são realizadas simulações do processamento local dos nós para determinar a redução no número total de amostras que cada nó sensor ativo deve enviar ao sorvedouro. A fração de amostras enviadas é utilizada como um índice de conservação de energia. Desta forma, preserva-se a generalidade dos resultados, evitando os efeitos da escolha de um MAC específico. As simulações realizadas baseiam-se no problema de medição de temperatura apresentado

na Seção 3. Estas simulações são feitas com o auxílio da ferramenta MATLAB e levam em consideração diversas possibilidades de configuração do mecanismo proposto. São variados três parâmetros principais: a frequência com que são enviadas atualizações sobre o comportamento esperado para os próximos dias, o fator  $\alpha$  utilizado na construção do comportamento esperado e o erro admissível por amostra.

A frequência com que o vetor de atualização do comportamento esperado é enviado tem influência direta sobre o número total de amostras enviadas. O número de amostras enviado nas atualizações é fixo em função da frequência de atualização. Quanto maior esta frequência, maior o número de amostras enviadas nas atualizações. No entanto, há uma tendência a aumentar o número de amostras enviadas por apresentarem diferenças em relação ao comportamento esperado maiores que o erro admissível na rede quando se aumenta muito o intervalo entre atualizações.

O fator  $\alpha$  pondera o comportamento esperado em relação às amostras de períodos mais recentes. Quanto maior o valor deste parâmetro, maior o peso que as amostras mais recentes apresentam no cálculo do comportamento esperado. Desta forma, no caso de haver uma mudança brusca no comportamento de um dia para o outro, o comportamento esperado será mais parecido com o comportamento dia imediatamente anterior, diminuindo o peso do comportamento histórico do processo.

O erro admissível na rede é um parâmetro que representa a qualidade requerida na reconstrução da informação enviada pelo sensor ao sorvedouro. Quanto menor este erro, melhor a estimação do campo sensoriado. No entanto, ao reduzir o erro permitido aumenta-se o número de amostras enviadas por apresentarem erro maior que o permitido e, assim, a redução no número de amostras enviadas pelo sensor ao sorvedouro será menor.

Um ponto de grande importância na análise deste mecanismo é a forma com que os dados de entrada no simulador são gerados. Estes dados sintéticos devem representar as principais características da aplicação focada que tenham maior influência nos resultados do esquema analisado [Yu et al., 2004]. A Seção 4.1 detalha o procedimento utilizado na geração dos dados utilizados para analisar o esquema proposto.

#### **4.1. Modelagem dos Dados de Entrada**

Os dados de entrada são modelados com base em informações obtidas a partir de dados de temperatura coletados por diferentes estações meteorológicas como os mostrados na Figura 2. A partir da observação dos dados coletados nas estações meteorológicas, pode-se perceber algumas características importantes na evolução desses dados ao longo de um mesmo dia e entre dias consecutivos. Com relação ao comportamento ao longo de um dia deve-se destacar dois pontos principais: a existência de um período no meio do dia onde a temperatura média eleva-se e a maior variação na temperatura observada em horários consecutivos de um mesmo dia nestes períodos de temperatura média mais elevada. Com relação ao comportamento em dias consecutivos pode-se perceber uma certa relação entre as temperaturas médias e as diferenças entre a maior e a menor temperatura em dias consecutivos.

É importante destacar que o objetivo na geração dos dados para analisar o esquema proposto não é reproduzir exatamente curvas de temperatura, mas reproduzir as características principais obtidas a partir da observação de dados reais. Desta forma, o processo empregado na geração de dados busca fornecer vetores com dados de entrada para as simulações que apresentem as características destacadas anteriormente. Isto permite uma melhor análise do desempenho do esquema proposto em função das principais características do processo sensoriado.

Primeiramente a base para a construção do comportamento da temperatura ao longo de cada dia foi um período da função  $-\cos(x)$ , de modo a obter a média mais alta de temperatura no meio do dia. Este função foi deslocada somando-se uma constante de modo a fazer com que a temperatura média comporte-se de forma parecida à mostrada na Figura 2(a).

Para modelar a relação existente entre as temperaturas médias de dias consecutivos, a constante somada à função  $-\cos(x)$  para alterar a temperatura média num dado dia  $i$  ( $const(i)$ ) é definida como a constante utilizada no dia  $i - 1$  ( $const(i - 1)$ ) somada a uma variável aleatória gaussiana com média 0 e desvio padrão  $\sigma_1$ . Quanto menor o valor de  $\sigma_1$ , maior a relação entre a temperatura média de um dia e a temperatura média do dia seguinte. O mesmo procedimento foi utilizado para modelar a amplitude da função  $-\cos(x)$  utilizada na geração dos dados do dia  $i$  ( $amp(i)$ ) em função da amplitude utilizada para a função  $-\cos(x)$  para o dia anterior  $amp(i - 1)$ . Esta amplitude representa a relação entre as diferenças da temperatura máxima para a temperatura mínima em dias consecutivos.

Por fim, a cada ponto da função  $-\cos(x)$  (que representam a temperatura em diferentes horários de um mesmo dia) foi somada uma variável aleatória com distribuição gaussiana também de média 0 mas com desvio padrão  $\sigma_2 \times C$ , onde  $C$  é proporcional ao valor da função  $-\cos(x)$  no horário considerado. Este comportamento de  $C$ , faz com que amostras de horários com temperaturas médias mais elevadas apresentem uma maior variação. Quanto maior  $\sigma_2$ , maior será esta variação. Todas as variáveis aleatórias gaussianas utilizadas tiveram seus valores máximos e mínimos limitados, evitando desta forma diferenças irreais de temperatura entre dias consecutivos. Desta forma, a variável gaussiana com seus valores máximos e mínimos limitados entre  $a$  e  $b$  ( $G_a^b$ ) possui uma função densidade de probabilidade igual a:

$$pdf(G_a^b) = \frac{f(x)}{\int_a^b f(x)} | a \leq x \leq b, \quad (3)$$

onde  $f(x)$  é a função densidade de probabilidade de uma gaussiana.

As  $k$  amostras geradas para um dado dia  $D_i$  são obtidas da seguinte forma:

$$D_i(k) = (amp(i - 1) \times G_a^b(0, \sigma_1)) \times \cos\left(\frac{2 \times k \times \pi}{N}\right) + \quad (4)$$

$$(const(i - 1) + G_a^b(0, \sigma_1)) + G_a^b(0, \sigma_2 \times C), \quad (5)$$

onde  $G_a^b(x, y)$  é uma variável gaussiana com média  $x$  e desvio padrão  $y$ , enquanto que  $N$  é o número de amostras obtidas para um mesmo dia.

Os dados diários são gerados com 96 amostras, representando a medição da temperatura a intervalos regulares de 15 minutos. Para uma melhor análise dos efeitos das características do processo sensoriado sobre o desempenho do esquema proposto, os valores de  $\sigma_1$  e  $\sigma_2$  são variados ao longo das simulações. Quanto maior estes dois desvios padrão, maior a imprevisibilidade do processo. A próxima seção apresenta maiores detalhes das simulações realizadas e os resultados obtidos.

## 4.2. Resultados

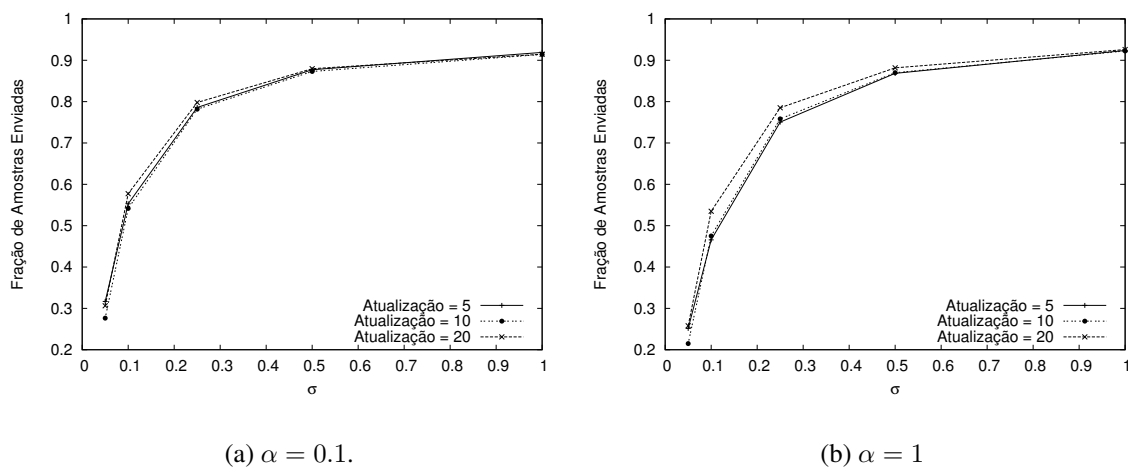
As simulações do processamento local dos nós são feitas utilizando dados gerados segundo a modelagem descrita na Seção 4.1 para representar os dados coletados por um nó sensor. O esquema proposto é aplicado a este conjunto de dados e calcula-se a fração do total de amostras que o nó sensor enviaria efetivamente ao sorvedouro. Quanto menor esta fração, melhor o desempenho do esquema proposto. A simulação assume que o nó



já identificou a periodicidade de 96 amostras do processo, de modo a analisar somente o funcionamento em longo prazo do esquema. Considerar o procedimento de identificação do período natural do processo geraria uma sobrecarga de amostras coletadas que poderia influir na análise dos resultados em regime permanente. Para eliminar este efeito, seria necessário alongar demasiadamente a simulação, tornando-a muito demorada. Além disso, este procedimento de identificação do período independe dos parâmetros de configuração do mecanismo. Somente as medidas coletadas pelo nó sensor influenciam neste procedimento.

O esquema proposto possui alguns parâmetros de configuração que podem influir nos seus resultados: a frequência de atualização, a constante  $\alpha$  e o erro admissível por amostra. Desta forma, as simulações são realizadas com a variação destes parâmetros para um melhor entendimento de seus efeitos. Em todas as simulações realizadas a frequência de atualização é um vetor de comportamento esperado enviado a cada *Atualização* dias. Desta forma, quanto maior o parâmetro *Atualização*, menor a frequência de atualização. O erro admissível por amostra é o parâmetro *erro máximo* multiplicado pelo comportamento esperado no horário em questão. Todos os resultados obtidos possuem confiabilidade de 99% e intervalos de confiança menores que 3% do valor médio.

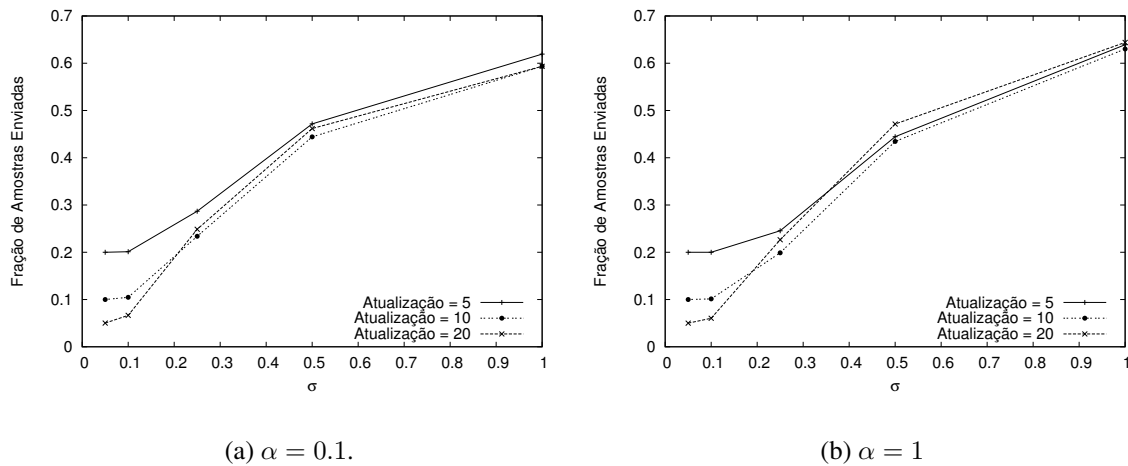
A primeira análise realizada é sobre a variação do número de amostras enviadas em função da incerteza no processo sensoriado. Para isso, é importante analisar o comportamento do número de amostras enviadas em função da variação de  $\sigma_1$  e  $\sigma_2$ . A Figura 4 mostra o número de amostras enviadas quando  $\sigma_1$  e  $\sigma_2$  variam juntos ( $\sigma$ ).



**Figura 4: Amostras enviadas com a variação de  $\sigma$  com erro máximo = 0.01.**

Analisando a Figura 4 pode-se perceber que para um valor muito baixo do erro admissível por amostra o parâmetro  $\alpha$  e a frequência de atualização não possuem grande influência no resultado. Deve-se destacar que a frequência mais baixa utilizada, *Atualização* = 20, apesar de enviar menos amostras quando a variação do processo é muito baixa, apresenta um desempenho destacadamente inferior às outras duas frequências consideradas à medida que a variação do processo aumenta com o incremento de  $\sigma$  ( $\sigma_1$  e  $\sigma_2$ ). Para este erro máximo admitido, a fração de amostras sobe rapidamente para 90% com o aumento de  $\sigma$ . A Figura 5 mostra a mesma análise com um erro admissível igual a 5% do valor esperado.

A partir da Figura 5, é possível notar que com o aumento do erro admissível, os parâmetros  $\alpha$  e *Atualização* passam a ter um peso maior nos resultados. Desta forma, a variação destes parâmetros causa mudanças mais perceptíveis nos resultados. Para um



**Figura 5: Amostras enviadas em função de  $\sigma$  e com erro máximo = 0.05.**

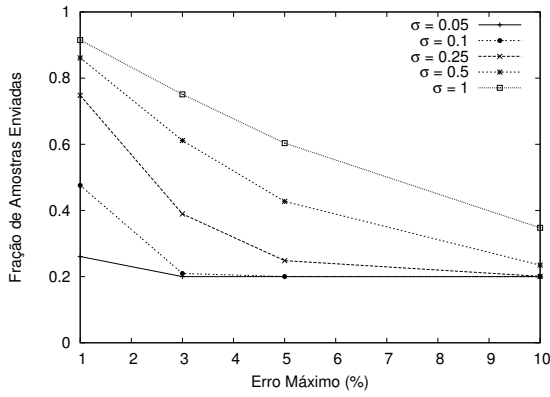
erro máximo de 5%, o esquema consegue reduzir em torno de 35 a 40% o número de amostras enviadas com  $\sigma = 1$ . Com a utilização de restrições muito fortes para o erro admissível, o ganho do esquema cai muito rapidamente, pois quanto maior a fração de amostras enviadas menor será o ganho do esquema. Além disso, é importante destacar que para valores mais elevados de  $\sigma$  a frequência de atualização de um comportamento esperado a cada 10 dias apresenta resultados melhores. Isto indica que o compromisso entre número fixo de amostras enviadas nas atualizações e o número de amostras enviadas por erro à medida que o intervalo entre atualizações aumenta é melhor atendido com esta frequência de atualização intermediária.

A influência mostrada pelo erro admissível no comportamento do esquema torna necessária uma análise mais detalhada dos resultados em função deste erro. A Figura 6 mostra os resultados em função do erro admissível. Nestes resultados é utilizado um valor intermediário do parâmetro  $\alpha$ .

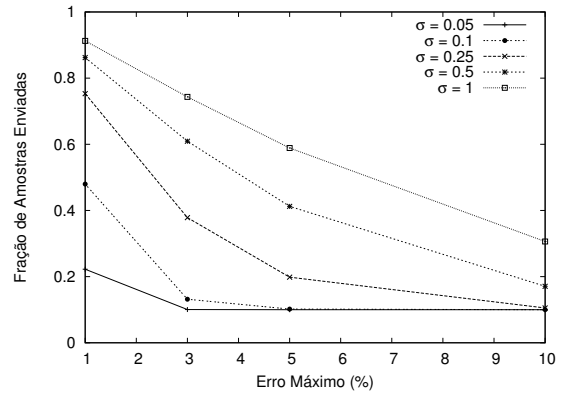
A Figura 6 mostra que independente dos valores de *Atualização* utilizados, à medida que o erro admissível diminui, as pequenas variações de  $\sigma$  ( $\sigma_1$  e  $\sigma_2$  juntos) pioram de forma mais acentuada o desempenho do esquema. Para erros admissíveis altos, pequenas variações de  $\sigma$  são negligenciáveis. Deve-se destacar que, para valores mais elevados do erro admissível e com valores baixos de  $\sigma$  (processos mais regulares), a fração de amostras enviadas aproxima-se da fração de amostras que deve ser enviada de forma fixa em função da frequência de atualização utilizada. Estas frações são 20, 10 e 5% para *Atualização* igual a 5, 10 e 20 dias, respectivamente. Além disso, mesmo para processos pouco regulares ( $\sigma = 1$ ), é possível obter uma redução em torno de 20 a 10% para erros admissíveis tão baixos quanto 3 e 1%.

A última análise realizada busca verificar o desempenho do mecanismo em função de  $\sigma_1$  e  $\sigma_2$  separadamente. Nesta análise, Figura 7, é utilizado um valor intermediário para o erro admissível de modo ao esquema não ser exageradamente sensível às variações de  $\sigma$ , como para o erro admissível igual a 1%, nem inatingido por estas variações, como para o erro admissível igual a 10%.

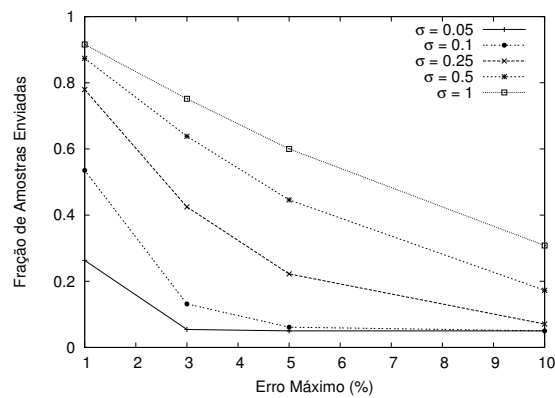
Enquanto que o esquema com  $\alpha = 1$  (Figura 7(a)) é quase que igualmente sensível à variação de  $\sigma_1$  e  $\sigma_2$ , quando o esquema utiliza  $\alpha = 0.1$  pode-se notar uma maior sensibilidade à variação de  $\sigma_1$ . A sensibilidade a  $\sigma_2$  é parecida nos dois casos ( $\alpha = 1$  e  $\alpha = 0.1$ ), pois  $\sigma_2$  não tem ligação direta com o fato de períodos consecutivos terem valores com algum tipo de relação. O parâmetro  $\sigma_1$ , por outro lado, regula a relação entre



(a) Atualização = 5.



(b) Atualização = 10.



(c) Atualização = 20.

**Figura 6: Amostras enviadas em função do erro máximo com  $\alpha = 0.25$ .**

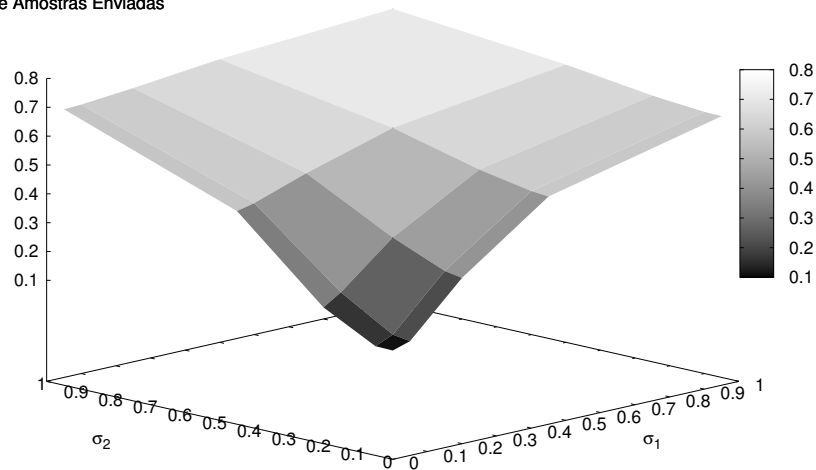
períodos consecutivos e o parâmetro  $\alpha$  busca adequar o desempenho do esquema a diferentes condições dessa relação entre amostras consecutivas.

## 5. Conclusões

Este artigo propõe e analisa um esquema bio-inspirado para reduzir o número de amostras enviadas por um nó sensor ao sorvedouro ao longo do tempo. Para isto, o esquema proposto explora as características específicas do processo físico sensoriado com o objetivo de identificar um comportamento regular para o processo e, desta forma, definir sua periferia de atenção. Eventos periféricos, ou seja, que se enquadrem na definição de periferia do nó, não são reportados para o sorvedouro. Este procedimento diminui o tráfego de dados na rede e, conseqüentemente, o consumo de energia dos nós sensores. O esquema proposto tem a vantagem de reduzir o envio de amostras mesmo dos nós que estejam em regiões de fronteira, o que algoritmos que buscam reduzir a densidade espacial de amostras enviadas não conseguem fazer. Além disso, o esquema proposto pode ser utilizado em conjunto com esses algoritmos para reduzir ainda mais o total de amostras enviadas ao nó sorvedouro.

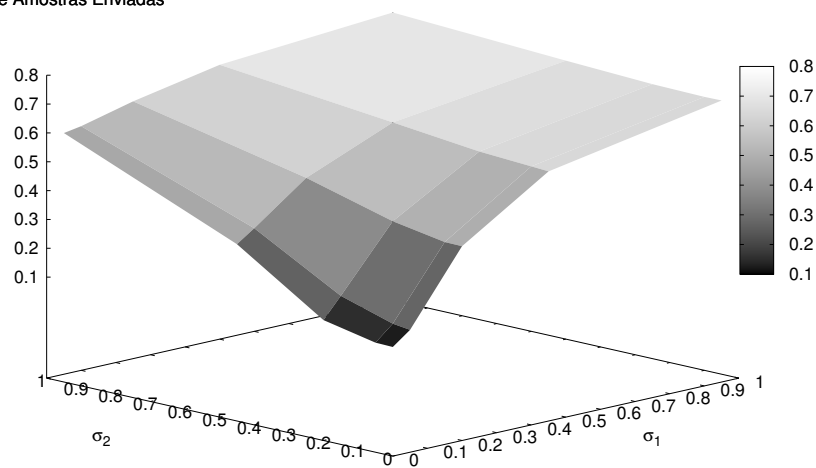
A viabilidade do esquema foi analisada numa aplicação modelada com base na medição de temperatura. A geração dos dados para simulação foi cuidadosamente desenvolvida de modo a reproduzir as principais características deste processo físico que podem influir no desempenho do esquema proposto. Os resultados mostram que existe um com-

Fração de Amostras Enviadas



(a)  $\alpha = 1$ .

Fração de Amostras Enviadas



(b)  $\alpha = 0.1$

**Figura 7: Amostras enviadas em função de  $\sigma_1$  e  $\sigma_2$  com erro máximo = 0.03 e atualização = 10.**

promisso entre o número de vezes que o nó sensor atualiza o comportamento esperado para o processo e o número de amostras enviadas por diferir muito deste comportamento. Restrições muito fortes para o erro admissível na reconstrução dos dados enviados pelos sensores ao sorvedouro podem reduzir rapidamente o ganho obtido com o esquema. Os dados utilizados como entrada para as análises foram gerados com diferentes graus de liberdade e os resultados obtidos sugerem que o esquema deve ser configurado levando em consideração o quanto um período do processo tem relação com o período anterior. Mesmo para processos pouco regulares, o esquema proposto apresenta reduções de 10 a

20% no total de amostras enviadas. Para processos com comportamentos mais regulares, o esquema pode enviar somente os vetores de atualização, que representam entre 5 e 20% do total de amostras, dependendo da frequência de atualização utilizada.

Como trabalhos futuros, está a análise do efeito de perdas na rede sobre a reconstrução do processo pelo nó sorvedouro, a utilização de dados coletados por sensores reais na avaliação do esquema e a análise dos efeitos da identificação da periodicidade do processo sensoriado sobre os resultados obtidos com o esquema proposto.

## Referências

- Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., and Cayirci, E. (2002). Wireless sensor networks: a survey. *Computer Networks*, 38:393–422.
- Batalin, M. A., Rahimi, M., Y. Yu, D. Liu, A. Kansal, Sukhatme, G., Kaiser, W., M. Hansen, Pottie, G. J., Srivastava, M., and Estrin, D. (2004). Towards event-aware adaptive sampling using static and mobile nodes. Technical Report 38, Center for Embedded Networked Sensing - CENS.
- Chen, H., Li, J., and Mohapatra, P. (2004). Race: Time series compression with rate adaptivity and error bound for sensor networks. In *IEEE International Conference on Mobile Ad-hoc and Sensor Systems - MASS 2004*.
- GEORIO (2003). *Alerta Rio*. Fundação Instituto de Geotécnica do Município do Rio de Janeiro. <http://www2.rio.rj.gov.br/georio/site/alerta/quadro.asp>.
- Kumar, A., Ishwar, P., and Ramchandran, K. (2004). On distributed sampling of smooth non-bandlimited fields. In *Information Processing In Sensor Networks - IPSN'04*, pages 89–98.
- Lazaridis, I. and Mehrotra, S. (2003). Capturing sensor-generated time series with quality guarantees. In *International Conference on Data Engineering (ICDE'03)*.
- Nowak, R., Mitra, U., and Willett, R. (2004). Estimating inhomogeneous fields using wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, 22(6):999–1006.
- Pottie, G. P. and Kaiser, W. J. (2000). Wireless integrated network sensors. *Communications of the ACM*, 43(5):51–58.
- Rahimi, M., Pon, R., Kaiser, W. J., Sukhatme, G. S., Estrin, D., and Srivastava, M. (2004). Adaptive sampling for environmental robotics. In *IEEE International Conference on Robotics & Automation*, pages 3537–3544.
- Weiser, M. and Brown, J. S. (1997). The coming age of calm technology. In *Beyond calculation: the next fifty years*, pages 75 – 85. Copernicus.
- Willett, R., Martin, A., and Nowak, R. (2004). Backcasting: adaptive sampling for sensor networks. In *Information Processing In Sensor Networks - IPSN'04*, pages 124 – 133.
- Yu, Y., Estrin, D., Rahimi, M., and Govindan, R. (2004). Using more realistic data models to evaluate sensor network data processing algorithms. In *IEEE Workshop on Embedded Networked Sensors - EmNetS-I*.