

Efeitos das Redes de Sensores sem Fios na Estimação de Parâmetros

Rian Gabriel Santos Pinheiro¹
Paulo Rogério de Souza e Silva Filho¹
Alejandro C. Frery^{1,2}

¹ Laboratório de Computação Científica e Análise Numérica – LaCCAN
Centro de Pesquisas em Matemática Computacional – CPMAT

² Laboratório de Computação Científica e Visualização – LCCV
Universidade Federal de Alagoas – UFAL

BR 104 Norte km 97 – 57072-970 – Maceió – AL, Brasil
rian.gabriel,progerio07@gmail.com, acfrery@pq.cnpq.br

Abstract. In this paper we analyze the behavior of estimators of parameters that characterize zero-mean isotropic gaussian random fields obeying the Matérn correlation model, when the fields are subjected to data sampling and fusion. These two operations characterize the operation of Wireless Sensor Networks, which are devices widely employed in the remote monitoring of environmental indicators. Using a Monte Carlo approach we conclude that these networks introduce large deviations on the estimates.

Keywords: simulation, signal reconstruction, geostatistics simulação, reconstrução de sinais, geoestatística

1. Redes de Sensores sem Fio

Dentre as diversas áreas de pesquisas que têm recebido atenção e destaque, as redes de sensores sem fio (RSSF) se inserem nesse grupo pela possibilidade de serem utilizadas em várias aplicações e pelos muitos desafios que elas oferecem e que ainda não foram solucionados plenamente. As redes de sensores sem fio hoje estão presentes em cenários de estudos diversos: automação e otimização de processos industriais, monitoramento de habitats e espécies em extinção, coleta de dados de ambientes inóspitos ou que ofereçam perigo aos seres humanos (vulcões, regiões de instabilidade sísmica, regiões suscetíveis a ocorrência de furacões etc), monitoramento das condições estruturais de construções antigas, monitoramento e coleta de informações sobre saúde de pacientes (*body area networks*), aplicações militares de rastreamento etc (LOUREIRO et al., 2003).

As RSSF são um subconjunto das redes *ad hoc*, entretanto possuem características próprias: são formadas por dezenas, centenas ou milhares de dispositivos sensores independentes (normalmente conhecidos por nós sensores ou simplesmente “nós”) e distribuídos espacialmente. Cada nó sensor é responsável por monitorar condições ambientais através da coleta de dados físicos diversos, tais como temperatura, pressão atmosférica, vibração, luminosidade e velocidade. Os nós realizam um pré-processamento e agregação dos dados obtidos (AQUINO; NAKAMURA, 2009), para depois encaminhá-los ao local de destino, que pode ser outro nó (denominado “nó *sink*”) ou a estação base à qual o usuário final tem acesso direto (AKYILDIZ et al., 2002; YICK; MUKHERJEE; GHOSAL, 2008).

Entretanto, as RSSF possuem várias limitações de processamento, armazenamento e consumo energético, sendo essas algumas das principais questões que motivam as contínuas pesquisas nessa área. Muitos estudos mostram que o consumo de energia das baterias dos nós sensores é bastante afetado pela transmissão de dados através do enlace sem fio (SILVA; FRERY; VIANA, 2010) e, portanto, diversas técnicas foram criadas e ainda estão em estudo para tratar da melhor forma o gasto energético, conforme demonstrado por Anastasi, Conti e Di Francesco

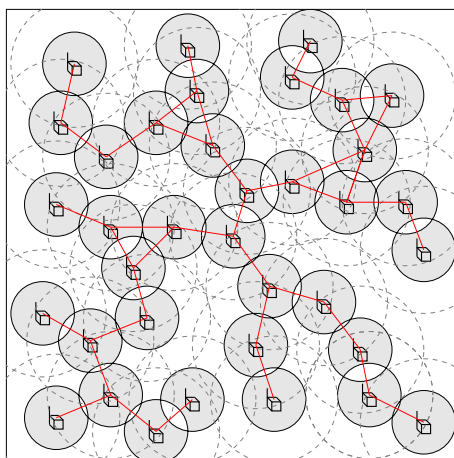


Figura 1: Exemplo de uma rede de sensores sem fio

(2009). Portanto, é bastante comum optar pelo uso de técnicas de agrupamento e fusão de dados, de forma a minimizar o gasto energético e maximizar o tempo de vida das redes de sensores.

Os dados obtidos pelo nó *sink* ou pela estação base podem ser utilizados para reconstruir informações a respeito de um fenômeno de interesse. Entretanto, essas técnicas têm impactos na qualidade do sinal reconstruído, que podem chegar a inviabilizar o seu uso. Frery et al. (2010) citam os principais fatores que impactam na qualidade do sinal reconstruído: a granularidade dos dados, a estratégia de amostragem, o processo de agrupamento dos nós sensores, a agregação e a reconstrução dos dados. A análise da qualidade do sinal reconstruído é um dos pontos levantados por Welsh (2010) como fundamentais para a pesquisa futura em RSSF.

Os impactos causados por essas técnicas podem ser encontrados em diversos trabalhos na literatura, como os de Nordio, Chiasserini e Muscariello (2010), Sung, Poor e Yu (2009). Todos os trabalhos anteriores presentes na literatura focam na reconstrução do sinal, entretanto, neste artigo analisamos a qualidade do sinal reconstruído através do modelo que pode ser inferido através dele, e não dos dados. Esta abordagem tem por objetivo analisar a degradação da informação em termos de modelos, e não de dados.

Na Seção 2 são descritos os modelos de agrupamento de sensores que foram utilizados, na Seção 3 abordamos o modelo utilizado para o campo aleatório gaussiano, enquanto na Seção 4 apresentamos o modelo de disposição dos sensores. Na Seção 5 apresentamos os detalhes da metodologia da pesquisa, seguidos das Seções 6 e 7, que abordam os resultados alcançados e as conclusões finais.

2. Modelos de agrupamento

Algoritmos hierárquicos de roteamento têm como objetivo agregar os sensores em *clusters* e revezar a escolha dos responsáveis pela entrega dos dados (ação que consome muita energia de cada sensor), proporcionando uma maior eficiência energética e, por consequência, maior duração do tempo de vida dessas redes. Dentre os mais difundidos na literatura, apresentaremos a seguir os algoritmos LEACH e SKATER.

2.1. LEACH

O LEACH (“*Low-Energy Adaptive Clustering Hierarchy*”) é um protocolo hierárquico de roteamento de dados que efetua agrupamentos de nós sensores de forma a otimizar o consumo energético da rede, garantindo assim uma maior autonomia da mesma (HEINZELMAN;

CHANDRAKASAN; BALAKRISHNAN, 2002). A cada rodada, cada nó utiliza um algoritmo estocástico para determinar a probabilidade dele se tornar o *cluster-head* (CH), ou seja, o nó responsável por receber os dados de grupo. Assim, em cada rodada, os *cluster-heads* atuais são responsáveis por fazer algum tipo de fusão dos dados dos demais nós e de enviar o resultado ao nó *sink* ou à estação base de processamento de dados.

O processo começa com a decisão da atribuição do papel de *cluster-head* ao nó. Caso o nó receba essa atribuição, ele fará um *broadcast* dessa informação e irá aguardar por respostas válidas oriundas de membros do cluster. Cada nó que não foi escolhido como CH envia uma mensagem de retorno, requisitando se agrupar ao *cluster-head* mais próximo (escolhido pela potência de sinal mais forte que chega até esse nó). Após cada sensor se agrupar ao seu respectivo *cluster-head*, os *cluster-heads* ficam responsáveis pela criação de uma tabela TDMA, que corresponde ao intervalo de tempo que cada nó terá para se comunicar e enviar seus dados ao seu *cluster-head*. Após criar e enviar esta tabela aos nós de seu grupo, inicia-se o monitoramento do ambiente.

Este protocolo pode ser considerado clássico da área de redes de sensores sem fios e é usado como referencial (*benchmark*) para a avaliação comparativa de outras estratégias.

2.2. SKATER

O SKATER (*Spatial 'K'luster Analysis by Tree Edge Removal* é uma abordagem que procura grupos com dados o mais correlacionados o possível (ASSUNÇÃO et al., 2006). O algoritmo utiliza como entrada um grafo não orientado conexo $\mathcal{G} = (V, L)$, em que os vértices V são os sensores e cada aresta (i, j) de L está presente se o sensor i poder se comunicar com o sensor j de acordo com o raio de comunicação dos sensores, com um peso proporcional à distancia (em algum espaço adequado) dos dados sensorizados entre os mesmos. Com esse pesos calcula-se a árvore geradora mínima de \mathcal{G} .

Com a árvore geradora mínima calculada, basta remover $\kappa - 1$ arestas, sendo κ é o número de *cluster* desejados, para obter κ subgrafos conexos. Para a escolha das arestas a serem removidas utiliza-se de um particionamento recursivo que prioriza a homogeneidade dos *clusters*.

Como o algoritmo não define a escolha do *cluster-head*, o mesmo foi escolhido uniformemente entre todos os membros do grupo.

3. Campo Gaussiano

Para simular o fenômeno foi utilizado um campo aleatório gaussiano estacionário isotrópico com correlação espacial descrita pelo modelo de Matérn.

O modelo de correlação Matérn é um dos mais utilizados na literatura (DIGGLE; RIBEIRO, 2007). Sua representação é dada na forma

$$\rho(u) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)} \left(\frac{u}{\phi}\right)^{\kappa} K_{\kappa}\left(\frac{u}{\phi}\right), \quad (1)$$

onde u é a distância entre variáveis aleatórias, K_{κ} é a função de Bessel de ordem κ , $\phi > 0$ é o parâmetro de escala, e $\kappa > 0$ é o parâmetro de suavização denominado "ordem". A Figura 2 mostra eventos deste modelo para $(\phi, \kappa) \in \{0.1, 0.5, 1, 5, 10\}^2$.

Neste trabalho estimaremos os parâmetros, κ e ϕ após o sinal original ser amostrado e agrupado pela RSSF.

4. Processos Pontuais

A distribuição dos sensores segue a proposta de Frery et al. (2008, 2010), que consiste em utilizar processos pontuais para tal. Diversos processos pontuais são capazes de descrever várias

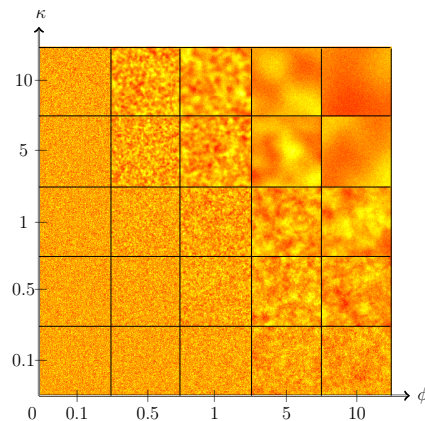


Figura 2: Variação de ϕ e κ em campos gaussianos seguindo o modelo de correlação Matérn

logísticas de localização de sensores no espaço de medição.

Neste trabalho consideramos que a distribuição espacial dos sensores segue o modelo SSI (*Simple Sequential Inhibition*) de Baddeley (2006) para distribuir os nós sensores em uma determinada região. Este modelo descreve a situação de controle parcial da distribuição de sensores, já que proíbe a ocorrência de sensores a uma distância inferior ao raio de inibição. Se o raio de inibição é nulo, tem-se um processo de total independência, ou Bernoulli.

Um evento do processo SSI de no máximo n pontos e raio de inibição $\alpha \geq 0$ numa região $R \subset \mathbb{R}^2$ pode ser construído iterativamente, estipulando um número máximo de iterações i_{\max} . O primeiro ponto é colocado segundo uma distribuição uniforme em R , e os subsequentes devem respeitar o raio de inibição, até os n pontos serem posicionados ou o número máximo de iterações i_{\max} ser alcançado. O Algoritmo 1 mostra um possível pseudocódigo (não otimizado) que implementa esta técnica.

Algoritmo 1 Simulação de um evento SSI

```

1: procedure SSI( $R, \alpha, n, i_{\max}$ )  ▷ Área do processo, raio de inibição, número máximo de
   partículas, número máximo de iterações
2:    $i \leftarrow 1$   ▷ Inicializa contador
3:   Posicionar a primeira partícula em  $(x_i, y_i)$  segundo uma lei uniforme em  $R$ .
4:   for  $2 \leq i \leq n$  do
5:      $\ell \leftarrow 0$ 
6:     repeat
7:        $T \leftarrow \text{TRUE}$ 
8:       Posicionar a partícula  $i$  uniformemente em  $R$ 
9:       for  $1 \leq j \leq i - 1$  do
10:         $T \leftarrow T \wedge [d((x_i, y_i), (x_j, y_j)) < \alpha]$   ▷ Verifica se a nova partícula está longe
   o suficiente de todas as outras
11:      end for
12:       $\ell \leftarrow \ell + 1$ 
13:    until  $T \vee \ell > i_{\max}$   ▷ Para ao colocar as  $n$  partículas ou ao atingir o número
   máximo de iterações
14:  end for
15:  return As coordenadas dos pontos
16: end procedure

```

5. Metodologia

Nesse trabalho foram utilizados os seguintes parâmetros: repulsão do processo SSI $\alpha \in \{0, 1, 4, 8\}$, ordem $\kappa \in \{0.1, 0.5, 1, 5, 10\}$ e escala $\phi \in \{0.1, 0.5, 1, 5, 10\}$. Com isso, analisamos um espaço paramétrico de dimensão $4 \times 5 \times 5 = 100$. Para cada ponto no espaço paramétrico foi escolhida uma semente única, e foram realizadas 1000 replicações, das quais só foram empregadas na análise aquelas que retornaram resultados viáveis, isto é, para o parâmetro θ , foram consideradas válidas replicações que retornassem estimadores dentro do intervalo $(\theta/10, 10\theta)$.

Foram realizadas simulações Monte Carlo com 100 sensores, cada um com raio de comunicação de 50 m, valor oriundo da especificação dos sensores MIKAZ da Crossbow (<http://www.xbow.com/>), e de percepção de 25 m, que redundava em uma boa amostragem do sinal com redundância aceitável. Apenas grafos conexos foram considerados. As simulações foram executadas no *cluster* GradeBR do Laboratório de Computação Científica e Visualização (LCCV), que possui 452 processadores Intel Nehalem, com 3 GB de memória RAM e 168 GB de armazenamento por processador.

Neste trabalho utilizamos a plataforma de análise estatística R (R Development Core Team, 2010) com os pacotes *spatstat* (BADDELEY; TURNER, 2005) para a geração da disposição dos sensores segundo o processo SSI, *spdep* (BIVAND et al., 2010) que contém a função *SKATER*, *RandomFields* (SCHLATHER, 2009) para gerar os campos gaussianos, *geoR* (RIBEIRO; DIGGLE, 2001) com uso da função *likfit* (máxima verossimilhança para campos gaussianos) para estimar os parâmetros e *lattice* (SARKAR, 2010) para a visualização dos dados multivariados.

6. Resultados

No que segue relatamos apenas os resultados da estimação da escala, por ser esse o parâmetro mais relevante do modelo. As Figuras 3, 4 e 5 mostram em linha vermelha o verdadeiro valor do parâmetro estimado em cada situação (cada caixa de cada figura). A primeira mostra resultados sem agrupamento, a segunda após o uso de LEACH e a terceira após o uso de SKATER.

Na Figura 3, que mostra os resultados sem agrupamento, percebemos que em todos os casos as estimativas ficam muito acima do valor verdadeiro. Há três situações para as quais não foi possível obter nenhuma replicação válida: $(\phi, \kappa) \in \{(5, 10), (1, 1), (1, 10)\}$.

Das Figuras 3, 4 e 5 concluímos que as estimativas do parâmetro ϕ ficam mais próximas do valor verdadeiro quando se utiliza *clusterização*, e que os resultados são ainda melhores para pequenos valores de ϕ . Concluímos ainda que, como utilizamos 100 sensores com raio de percepção de 25 m em um espaço de 100 m², o parâmetro α do processo SSI interfere pouco na estimação.

Em todos os casos analisados, o estimador de ϕ mostrou-se viesado.

7. Conclusão

As RSSF introduzem grandes desvios nas estimativas dos parâmetros que caracterizam o modelo do sinal, quando este é descrito por um campo gaussiano isotrópico com correlação Matérn.

Uma análise quantitativa dos resultados será reportada em outro trabalho. Em função do viés observado serão propostas técnicas para melhorar a estimação e, com isso, aprimorar o uso de inferência em RSSF.

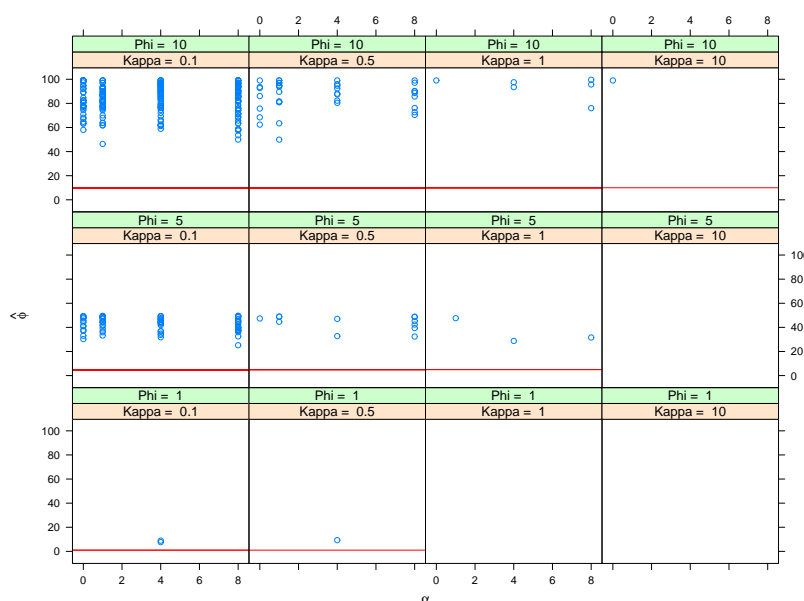


Figura 3: $\hat{\phi}$ estimado em redes sem agrupamento

8. Agradecimentos

Agradecemos a Bruno Lopes Vieira pelas orientações e esclarecimentos sobre a realização das simulações, ao prof. André L. L. Aquino e a Alyson L. C. Oliveira os comentários e questionamentos que refinaram esta pesquisa. Agradecemos a ANP, a Petrobrás e ao Laboratório de Computação Científica e Visualização da Universidade Federal de Alagoas por garantir acesso aos recursos computacionais do cluster GradeBR/UFAL da Rede Galileu. CNPq e Fapeal financiaram parcialmente esta pesquisa.

Referências

AKYILDIZ, I. F. et al. Wireless sensor networks: a survey. *Computer Networks*, v. 38, n. 4, p. 393–422, março 2002.

ANASTASI, G.; CONTI, M.; Di Francesco, M. Extending the lifetime of wireless sensor networks through adaptative sleep. *IEEE Transactions on Industrial Informatics*, v. 5, n. 3, p. 351 – 365, agosto 2009.

AQUINO, A. L. L.; NAKAMURA, E. F. Data centric sensor stream reduction for real-time applications in wireless sensor networks. *Sensors*, v. 9, p. 9666–9688, 2009.

ASSUNÇÃO, R. M. et al. Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, v. 20, n. 7, p. 797 – 811, August 2006.

BADDELEY, A. Spatial point process and their applications. In: *Lecture Notes in Mathematics*. [S.l.: s.n.], 2006. (Stochastic Geometry, v. 1892), p. 1–75. ISSN 1617-9692.

BADDELEY, A.; TURNER, R. Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, v. 12, n. 6, p. 1–42, 2005. ISSN 1548-7660. Disponível em: <www.jstatsoft.org>.

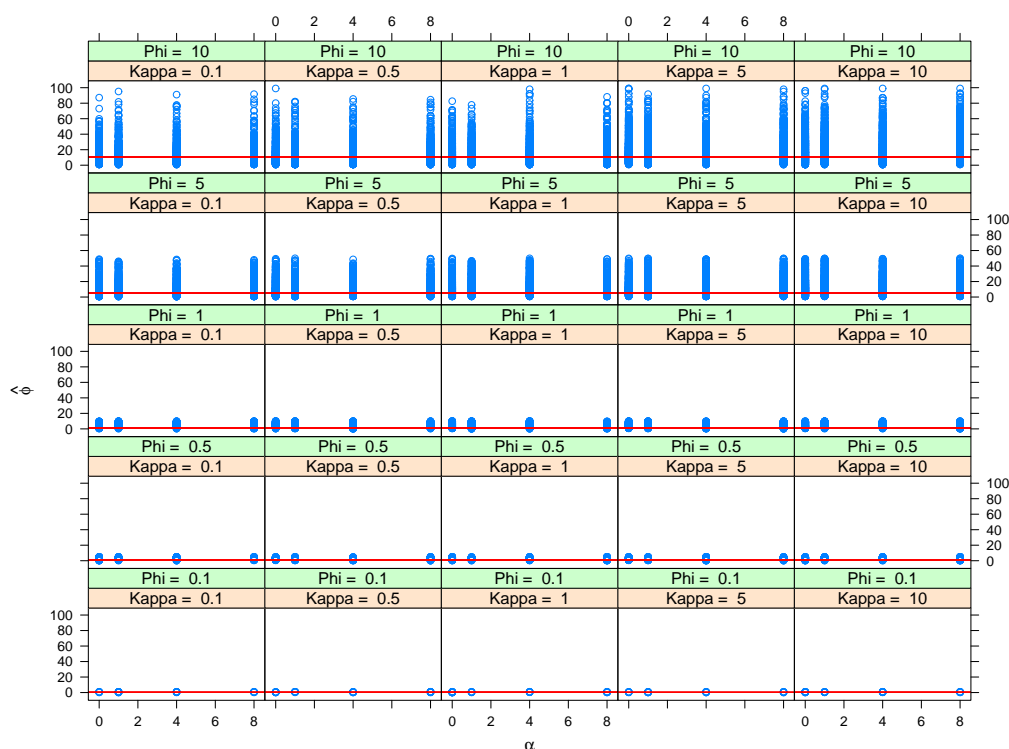


Figura 4: $\hat{\phi}$ estimado de uma rede com protocolo LEACH

BIVAND, R. et al. *spdep: Spatial dependence: weighting schemes, statistics and models*. [S.l.], 2010. R package version 0.5-19. Disponível em: <<http://cran.r-project.org/package=spdep>>.

DIGGLE, P. J.; RIBEIRO, P. J. *Model-based Geostatistics*. [S.l.]: Springer, 2007. (Springer Series in Statistics). ISBN 978-0-387-32907-9.

FRERY, A. C. et al. Error estimation in wireless sensor networks. In: ASSOCIATION FOR COMPUTING MACHINERY - ACM. *The 23rd Annual ACM Symposium on Applied Computing*. [S.l.], 2008. p. 1923–1928.

FRERY, A. C. et al. Data driven performance evaluation of wireless sensor networks. *Sensors*, v. 10, p. 2150–2168, 2010.

HEINZELMAN, W. B.; CHANDRAKASAN, A. P.; BALAKRISHNAN, H. An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, v. 1, n. 4, p. 660–670, outubro 2002.

LOUREIRO, A. A. F. et al. Redes de sensores sem fio. In: *Anais do XXI Simpósio Brasileiro de Redes de Computadores*. [S.l.: s.n.], 2003. p. 179–226.

NORDIO, A.; CHIASSERINI, C.; MUSCARIELLO, A. Signal reconstruction in sensor networks with flat and clustered topologies. *Computer Networks*, Elsevier North-Holland, Inc., New York, NY, USA, v. 54, p. 991–1004, April 2010. ISSN 1389-1286. Disponível em: <<http://dx.doi.org/10.1016/j.comnet.2009.09.023>>.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2010. ISBN 3-900051-07-0. Disponível em: <<http://www.r-project.org>>.

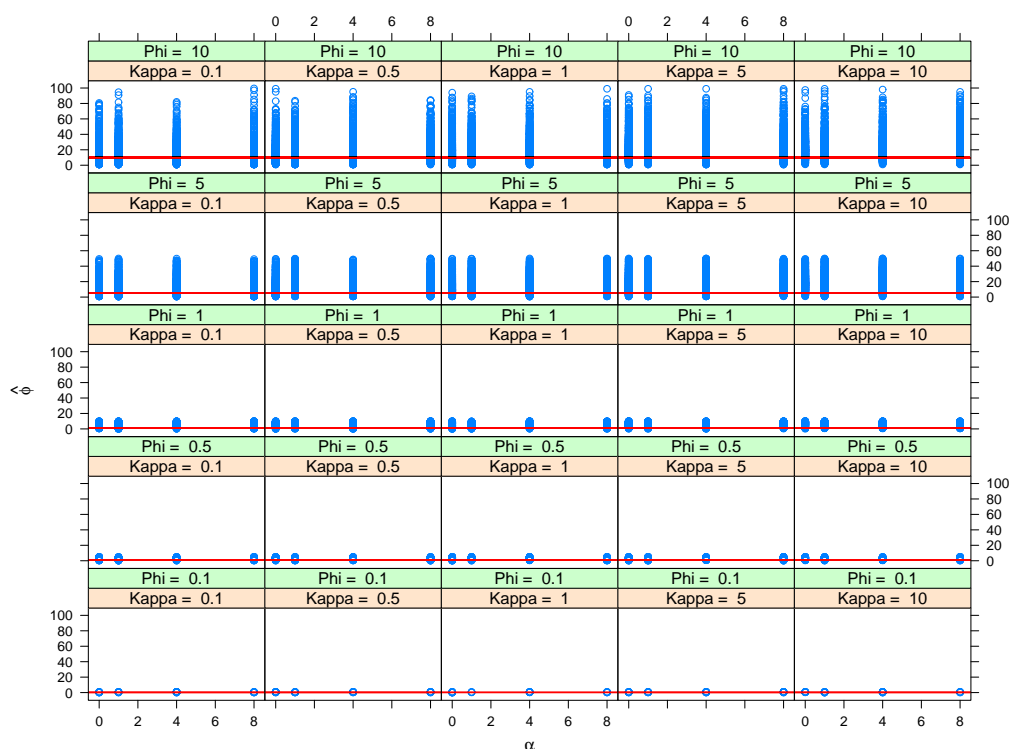


Figura 5: $\hat{\phi}$ estimado de uma rede com protocolo SKATER

RIBEIRO, P. J.; DIGGLE, P. J. *geoR: a package for geostatistical analysis*. *R-NEWS*, v. 1, n. 2, p. 14–18, June 2001. ISSN 1609-3631. Disponível em: <<http://cran.r-project.org/doc/Rnews>>.

SARKAR, D. *lattice: Lattice Graphics*. [S.l.], 2010. R package version 0.18-8. Disponível em: <<http://cran.r-project.org/package=lattice>>.

SCHLATHER, M. *RandomFields: Simulation and Analysis of Random Fields*. [S.l.], 2009. R package version 1.3.41. Disponível em: <<http://cran.r-project.org/package=RandomFields>>.

SILVA, C. V.; FRERY, A. C.; VIANA, P. Avaliação do consumo de energia e tempo de vida de redes de sensores sem fio comerciais. In: DVORKIN, E.; GOLDSCHMIT, M.; STORTI, M. (Ed.). *Mecânica Computacional*. Buenos Aires, Argentina, 2010. XXIX, p. 2663–2672.

SUNG, Y.; POOR, H. V.; YU, H. How much information can one get from a wireless ad hoc sensor network over a correlated random field? *IEEE Trans. Inf. Theor.*, IEEE Press, Piscataway, NJ, USA, v. 55, p. 2827–2847, June 2009. ISSN 0018-9448. Disponível em: <<http://portal.acm.org/citation.cfm?id=1669524.1669555>>.

WELSH, M. Sensor networks for the sciences. *Communications of the ACM*, v. 53, n. 11, p. 36–39, 2010.

YICK, J.; MUKHERJEE, B.; GHOSAL, D. Wireless sensor network survey. *Computer Networks*, v. 52, p. 2292–2330, 2008.