

Um Algoritmo Emergente para Coleta de Dados em Redes de Sensores sem Fio

Otávio Alcântara de Lima Júnior¹, Helano S. Castro² e Paulo Cesar Cortez²

¹Departamento de Telemática - Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)
Campus Maracanaú – Av. Contorno Norte, 10 – CEP: 61925-315 – Maracanaú – CE – Brasil

²Departamento de Engenharia de Teleinformática – Universidade Federal do Ceará (UFC)
Campus do Pici – CEP:60455-970 – Fortaleza – CE – Brasil

otavio@ifce.edu.br, helano@lesc.ufc.br, cortez@lesc.ufc.br

Abstract. *Many natural and man-made systems present emergent properties, where simple and repeated interactions among internal components lead to complex global behavior patterns. This emergent behavior can be explored by designers to reach application goals. This paper presents a wireless sensor network data gathering algorithm, which defines a set of simple rules to the nodes, achieving systemic behavior. That reduces messages exchanges and increasing network lifetime by reducing power consumption. Tests experiments demonstrated that the algorithm has a higher performance than similar techniques, in scenarios which the nodes has a failure probability.*

Resumo. *Vários sistemas naturais e criados pelo homem apresentam características emergentes, onde as interações simples e repetidas dos componentes internos geram padrões complexos de comportamento em nível global. Esse comportamento emergente pode ser explorado em nível de projeto para alcançar os objetivos da aplicação. Este trabalho apresenta uma proposta de algoritmo para coleta de dados em redes de sensores sem fio que define uma série de regras simples aos nós, obtendo um comportamento sistêmico, que reduz o número de mensagens trocadas pelos nós. Os experimentos realizados foram capazes de demonstrar que este algoritmo possui um desempenho aceitável, em relação a técnicas similares, em cenários onde os nós possuem probabilidade de falhar.*

1. Introdução

Redes de Sensores sem Fio (RSSFs) consistem de pequenos nós sensores inseridos no ambiente, e que tem a missão de extrair dados de monitoramento e enviá-los para análise. Os nós de uma RSSF são geralmente limitados pela potência de sua fonte de alimentação, poder computacional e alcance de seus enlaces de rádio. As RSSFs são utilizadas em diversos segmentos como monitoramento de estruturas, ambientes militares, dentre outros [Culler et al. 2004].

A arquitetura básica de uma RSSF está ilustrada na Figura 1. Os nós podem ser organizados em grupos ou atuarem individualmente, e os dados coletados são enviados para o nó sorvedouro que os disponibiliza para outras redes de comunicação. O roteamento das mensagens é centrado nos dados e não nos endereços. Nessas aplicações, é mais importante saber os valores ou eventos de uma determinada região do que o valor

lido por um sensor específico, sendo que a alta densidade de nós permite que o mesmo evento seja monitorado por diversos nós, criando uma redundância nas informações e maior tolerância a falhas.

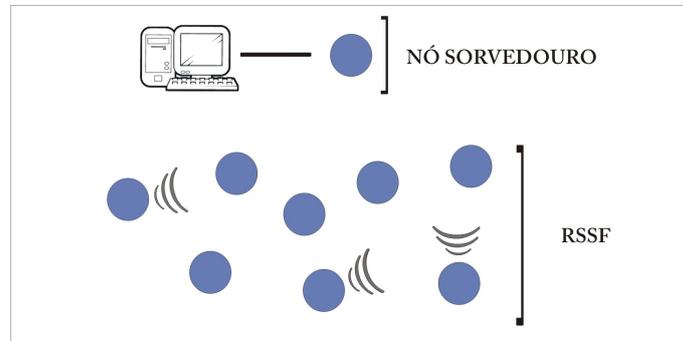


Figura 1. Modelo de Organização de uma RSSF típica.

O paradigma de comunicação das RSSFs diverge do modelo empregado em outras redes sem fio. Em uma RSSF, os nós da rede são instalados geralmente, em regiões de difícil acesso que inviabilizam a manutenção dos mesmos, bem como a troca de baterias. Por isso, questões de otimização do consumo de energia norteiam as decisões de projeto, que devem priorizar o prolongamento do tempo de vida da rede. Além disso, o fluxo de comunicação na rede é de muitos nós para um, e os nós sensores tendem a enviar os resultados de suas amostragens para o nó sorvedouro, numa espécie de multicast inverso. Por causa dessas e de outras especificidades, diversos protocolos de comunicação foram propostos para as RSSFs [Akkaya and Younis 2005].

Sabe-se que o consumo de energia é uma questão central no projeto de RSSFs, cuja principal técnica utilizada para reduzir o consumo é a fusão de dados. Esta técnica processa dados provenientes de diversas fontes e gera uma saída que é, de alguma forma, melhor do que as saídas das fontes individuais, através da cooperação entre os nós sensores. Dessa maneira, é possível reduzir o número de pacotes que trafegam pela rede aumentando seu tempo de vida [Nakamura et al. 2007].

As RSSFs são exemplos de sistema embarcado e distribuído. Nestas redes, há um fenômeno natural recorrente a auto-organização, indicando que a interação entre os componentes no nível microscópico acarretam comportamentos globais que emergem no nível macroscópico [Mills 2007]. Esses comportamentos emergentes podem ser indesejáveis e prejudiciais ao sistema, porém o projetista pode, intencionalmente, adicionar regras locais nos componentes que gerem comportamentos emergentes benéficos para a aplicação. Mills [Mills 2007] apresenta diversas abordagens que estimulam a auto-organização em RSSFs com o intuito de alocar banda de comunicação, formar grupos de sensores, disseminar informações, e organizar tarefas.

Este trabalho apresenta uma proposta de algoritmo para coleta de dados em RSSFs que utiliza os conceitos de sistemas auto-organizáveis, como forma de reduzir o número de mensagens enviadas e prolongar o tempo de vida da rede através da diminuição do consumo de energia.

O restante do trabalho está dividido da seguinte forma: a seção dois apresenta outros trabalhos relacionados ao tema; na seção seguinte são discutidas as principais

características de sistemas auto-organizáveis; a secção quatro apresenta o algoritmo proposto; em seguida apresentamos os resultados obtidos com os experimentos e a última secção trata as considerações finais.

2. Trabalhos Relacionados

Diversos protocolos para RSSFs se apóiam na interação local dos nós para atingirem objetivos globais, dentre estes podemos destacar a difusão direcionada [Intanagonwiwat et al. 2000]. Usando esse paradigma de comunicação, os dados gerados pelos sensores são nomeados através de pares de atributo-valor. Um nó solicita dados enviando uma mensagem de interesse especificando o tipo de evento desejado e a região de monitoramento. Quando um sensor da região de monitoramento recebe o interesse, este sensor passa a monitorar o ambiente buscando identificar o evento descrito. No momento que o evento é identificado, a informação retorna pelo caminho inverso da propagação do interesse. Os nós intermediários podem agregar novos dados, melhorando a precisão das medições. Essas interações são realizadas utilizando apenas troca de mensagens entre nós vizinhos. Contudo, esse algoritmo não se adequa bem para aplicações de monitoramento contínuo de ambientes.

Outro protocolo de RSSFs que busca inspiração no comportamento de sistemas biológicos foi apresentado por Cunha(2005). O objetivo desse protocolo é fornecer uma estimativa de processos variáveis espaço-temporalmente em uma dada região. Cada nó da rede identifica padrões no processo físico e reporta ao sorvedouro apenas os comportamentos que desviem do esperado [Cunha and Duarte 2005]. Por isso, esse protocolo só pode ser utilizado para monitorar processos físicos que apresentem um comportamento ao longo do tempo com alguma característica que possa ser considerada comum, ou esperada.

Madden(2002) apresenta uma RSSF, cujos nós são organizados em uma árvore de agregação. O nó raiz da árvore é o nó que envia as requisições e recebe seus resultados agregados. A requisição/resposta é propagada dos nós pais para os nós filhos. Os nós folhas enviam suas próprias medições para o pai. Nós intermediários esperam pelos valores de seus nós filhos, realizam a fusão de dados locais desses valores com suas próprias medições e envia o resultado para o seu nó pai. O nó raiz computa da mesma maneira que os nós intermediários e apresenta o resultado da agregação dos dados para o usuário [Madden et al. 2002]. Esse algoritmo requer a construção e manutenção de uma árvore de extensão. Em um ambiente real, falhas nos nós intermediários exigiriam a rotineira reconstrução da árvore de extensão. A Figura 2 ilustra a organização da rede nesse algoritmo.

Outra técnica de agregação de dados em RSSF é o protocolo *Push-Sum* [Kempe et al. 2003]. Cada nó mantém duas quantidades: um peso e um agregado. Em cada rodada, cada nó escolhe aleatoriamente um nó na rede e envia metade de seu peso e de seu agregado e o nó receptor adiciona essas quantidades as suas próprias. Se cada nó da rede for capaz de, aleatoriamente, contactar qualquer outro nó, o algoritmo calcula o valor agregado das medições da rede em um tempo que é proporcional ao logaritmo do número de nós. O cálculo do valor agregado é realizado de forma completamente descentralizada, mas para redes maiores o tempo de convergência pode se tornar um empecilho.

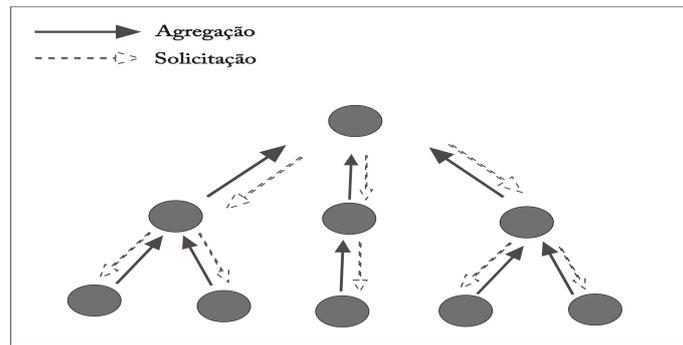


Figura 2. Estrutura da Árvore de Agregação.

3. Princípios de Auto-Organização

Um sistema com diversos componentes pode exibir comportamentos globais que aparentam ser mais complexos do que o comportamento de suas partes. Estes comportamentos emergentes aparecem naturalmente no processo de auto-organização do sistema, que é comum em vários sistemas naturais e criados pelo homem [Mills 2007]. Um exemplo são as propriedades de temperatura e pressão dos gases, que é uma medida da força das interações entre as moléculas. O projetista de protocolos para RSSFs pode tirar proveito dessas propriedades emergentes para atacar diferentes questões, como: alocação de banda, formação de grupos, fusão de dados, otimização do consumo de energia, dentre outras. No entanto, para tal é necessário ter o entendimento de quais princípios guiam e definem o comportamento emergente desses sistemas.

Sumpter (2006) enumera os princípios básicos inerentes a diferentes sistemas auto-organizáveis que usamos para entender os algoritmos que exploram o comportamento emergente. Os princípios são: integridade e variabilidade, reforço positivo, reforço negativo, limites de resposta, liderança, redundância, sincronização e egoísmo. Muitas das grandes questões do futuro dessa área se preocuparam em como trabalhar com esses princípios para gerar padrões de comportamento coletivo complexos. A seguir, discutiremos brevemente cada princípio.

Integridade e variabilidade estão relacionadas à individualidade de cada componente do sistema. É importante que cada componente possua um grau de diferenciação dos outros, é uma forma de prover novas soluções para os problemas que o grupo deseja resolver. O reforço positivo é uma forma de um indivíduo compartilhar informações com o grupo, influenciando o comportamento dos outros. Por outro lado, o reforço negativo dissuade um indivíduo a participar de uma ação, permitindo criar estabilidade no padrão do grupo. Limites de resposta estão atrelados a quantidade de estímulos necessária para um indivíduo gerar uma resposta ao sistema.

Embora liderança possa parecer pouco associável com sistemas auto-organizáveis, existem diversos exemplos de sistemas nos quais alguns indivíduos são chave para organizar as ações do grupo. Redundância é outra característica presente em sistemas auto-organizáveis, pois, permite que o sistema continue funcionando mesmo que haja uma redução no número de componentes. Sincronização é um exemplo do reforço positivo, que contribui para a melhor coordenação das ações dos componentes.

Por fim, o egoísmo que define que os atos altruístas de um indivíduo devem possuir

relação custo/benefício menor do que o relacionamento com o indivíduo que lucra com esse ato [Sumpter 2006]. Nesse trabalho, focamos na utilização de três dos princípios listados acima: integridade e variabilidade, reforço positivo e reforço negativo. Esses princípios nortearam a definição das regras impostas aos nós sensores.

4. Modelo do Algoritmo Proposto

Neste trabalho, utilizar-se agregação de dados em árvore para obter uma estimativa de um processo físico variável espaço-temporalmente em uma dada região. O algoritmo é baseado na interação entre nós vizinhos, indicando os nós que possuem medições divergentes em relação a uma média das medições. Apenas os nós localizados nas regiões críticas, cujas medições divergem significativamente da média da rede, reportam suas medições, que são agregadas ao longo do caminho de transmissão ao nó sorvedouro, que poderá, por sua vez, alterar os parâmetros que definem a classificação dos nós. Desta forma, o nó sorvedouro possui uma visão geral da variação do processo físico monitorado.

Antes de apresentarmos a estrutura do algoritmo proposto, devemos especificar as características que uma RSSF deve possuir para sua implantação. Os nós da rede devem ser fixos, cientes de sua localização geográfica e possuir *hardware* semelhante. A densidade dos nós deve ser adequada para gerar correlação nas medições e cobrir toda a área de monitoramento. A demanda de dados pela aplicação deve ser relativamente baixa, como também os requisitos de latência. De um modo geral, o usuário deve estar interessado em ter uma estimativa da variação espaço-temporal de um processo físico em determinada região.

O algoritmo define que os nós da rede devem compartilhar o mesmo ciclo de vida, que se resume a: enviar uma mensagem de *broadcast* com informações sobre o próprio nó e sua vizinhança; esperar um tempo para receber as mensagens de seus vizinhos e depois desligar o rádio e o processador até o próximo período. A partir dessa mensagem periódica, os nós são capazes de realizar algumas atividades complexas, como: sincronização; agrupamento de nós; agregação de dados; difusão das solicitações e respostas a solicitações.

A primeira etapa do algoritmo é a sincronização, que é baseada na técnica de sincronização *firefly* [Breza and McCan 2008]. Uma mensagem periódica deve conter um campo que indique qual é o tempo para a próxima ativação do nó. Os nós devem ajustar o período de ativação pela média dos períodos dos nós da vizinhança. Assim, em algumas iterações todos os nós da rede devem estar sincronizados. Essa técnica permite que os nós deixem os seus rádios desligados a maior parte do tempo, prolongando o tempo de vida da rede.

O próximo passo do algoritmo é o agrupamento dos nós em regiões, sendo que outro campo existente na mensagem identifica a qual região o nó pertence. O nó sorvedouro sempre pertence à região zero. Os nós que estão ao alcance de seu rádio, ao receberem a mensagem, configura sua região para nível um. De forma similar, os outros nós da rede configuram sua região como sendo uma região de número acima do nó vizinho de menor região. Dessa forma, em poucas iterações, todos os nós da rede conseguem determinar suas regiões. A Figura 3 ilustra a topologia da rede, após os nós identificarem suas regiões, que refletem sua distância em relação ao nó sorvedouro.

O foco do algoritmo é gerar uma representação média das medições da rede, que

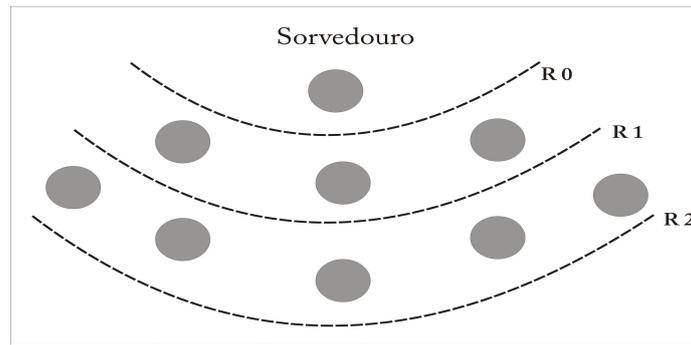


Figura 3. Regiões da RSSF.

é alcançada através da geração de uma estrutura de agregação em árvore. A cada nova iteração do algoritmo, cada nó da rede seleciona aleatoriamente um nó vizinho, que esteja na região imediatamente inferior. Este então envia para esse nó as suas medições, e o nó receptor agrega suas medições com as medições recebidas e as envia na próxima iteração. O valor agregado é representado por um par de números reais, o primeiro valor representa a soma das medições recebidas durante a iteração e o segundo representa o número de nós que enviaram suas medições. De forma que, em algumas iterações, as medições das sub-árvores chegam ao nó sorvedouro, que é capaz de calcular o valor médio de toda a rede. A Figura 4 ilustra esse algoritmo e é importante notar que essas informações são trocadas na mensagem periódica de sincronização dos nós.

Algoritmo

- 0 - Seja i um nó da rede alocado na região R ;
- 1 - Seja $\{w_k\}$ todos os valores enviados para o nó i pelos k nós de sua sub-árvore na iteração t ;
- 2 - Seja $W_{t,i} = \{\sum_j w_j, k\}$, o valor agregado calculado pelo nó i na iteração t ;
- 3 - Escolha aleatoriamente um nó da região $R-1$, e envie $W_{z,i}$ na iteração $t + 1$.

Figura 4. Algoritmo de Agregação.

Após calcular a média da rede, o nó sorvedouro adiciona esse valor na mensagem da próxima iteração, com o intuito de que todos os nós da rede tenham ciência do valor calculado. Assim, os nós podem utilizar esse valor para realizarem suas classificações, utilizando os parâmetros que são difundidos na rede pelo nó sorvedouro.

Neste ponto do algoritmo, os nós estão sincronizados, agrupados em regiões e cientes do valor médio das medições da rede. Para que o usuário obtenha uma visão mais precisa dos eventos da rede, é necessário que haja um mecanismo de difusão de solicitações de dados, que são os parâmetros utilizados pelo nó para identificar se está em

uma região de interesse de monitoração. Neste algoritmo, nós propomos utilizar a própria mensagem de sincronização para difundir as solicitações, que neste caso são bem simples. Um exemplo de solicitação de dados que foi implementada é que o nó verifique se sua amostragem local está dentro de um limite percentual em relação à média da rede.

As respostas das solicitações são enviadas pelo mesmo mecanismo de árvore de roteamento utilizado para agregar os valores das medições dos sensores, utilizando agregação de dados sempre que for necessário ao longo da rota. Contudo, apenas os nós que satisfazem os critérios da solicitação encaminham seus dados.

5. Simulação e Resultados

A simulação para a obtenção dos resultados apresentados neste trabalho se baseia no modelo escrito para o simulador SHOX [Lessman et al. 2008], onde os nós foram posicionados de forma aleatória em uma região quadrada de 50 m por 50 m. O raio de alcance dos rádios foi configurado como 20 m.

Os resultados apresentados se referem a um número de sensores variando entre 20 e 50. Os nós trocam mensagens a cada um segundo, em que 80% desse tempo é passado com o rádio desligado.

São realizados três tipos de experimentos para avaliar o desempenho do algoritmo proposto:

- verificação do tempo para que os nós entrem em sincronismo;
- verificação do tempo para completar o cálculo da agregação;
- cálculo do erro percentual entre o valor real da média e o calculado pelo algoritmo, para cenários onde os nós tem uma probabilidade p de serem desativados em cada iteração. As medições foram realizadas para três casos:
 - o nó agregador é escolhido aleatoriamente entre os nós da região imediatamente inferior, que foi a proposta de nosso algoritmo;
 - o nó agregador é o nó mais próximo que esteja na região imediatamente inferior, e não há cache dos valores agregados em iterações anteriores;
 - o nó agregador é o nó mais próximo que esteja na região imediatamente inferior, e os nós armazenam os valores agregados anteriores por quatro iterações.

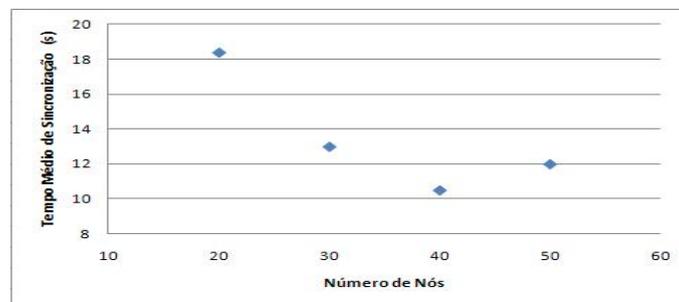


Figura 5. Tempo para Sincronização

Para cada caso de teste proposto nos experimentos foram realizadas um número arbitrário de cem iterações do algoritmo. As Figuras 5 e 6 mostram o tempo necessário

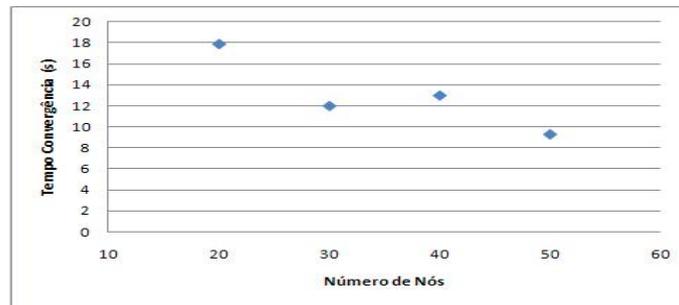


Figura 6. Tempo de Convergência da Agregação

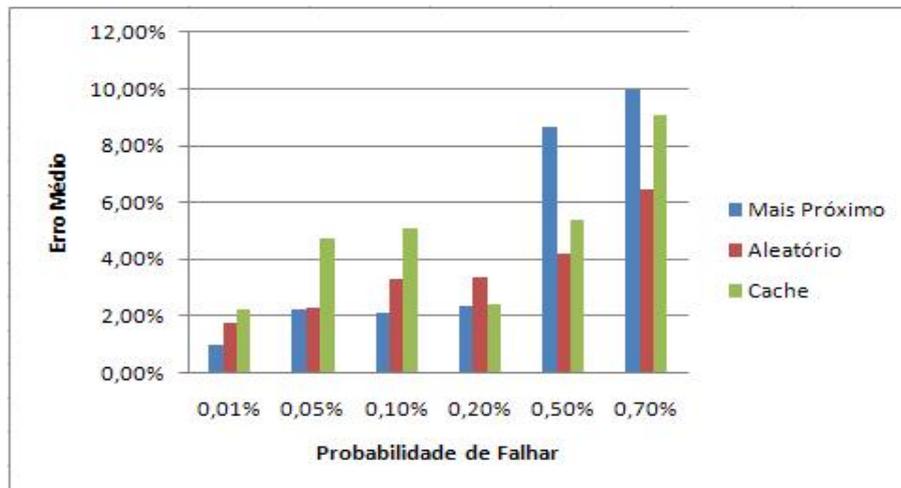


Figura 7. Erro Médio Verificado

para a convergência do sincronismo e do valor de agregado calculado pela rede. Como é de se esperar para redes menos densas, o tempo de convergência tende a ser maior.

O erro médio percentual entre o valor da agregação e o valor real da média da rede está mostrado na Figura 7. Nesse experimento, os nós possuem probabilidade p de falhar a cada iteração. Para cenários onde os nós tem alta probabilidade de falhar, escolher aleatoriamente o nó que fará parte da árvore de agregação produz um desempenho melhor do que escolher o nó mais próximo ou fazer cache de dados antigos. Para cenários com baixa probabilidade de falhas, o desempenho dos algoritmos é similar. É interessante notar que, para o cenário modelado, fazer cache de dados de agregação antigos não traz um bom desempenho devido ao fato do algoritmo de agregação usar dados de nós sensores que já estariam desativados.

6. Considerações Finais

A utilização dos conceitos de sistemas auto-organizáveis em projetos de sistemas distribuídos, em especial de RSSFs é importante e pode dar mais contribuições para a construção de protocolos que tirem maior proveito das capacidades desses sistemas. Neste artigo, utilizamos os princípios de auto-organização em RSSF para criar um algoritmo emergente que fornece uma descrição resumida da variação de um processo físico em uma determinada região. Os nós da rede trocam uma única mensagem por iteração do algoritmo que permite: sincronizar os nós; calcular a média das medições da rede e agregar

as respostas das solicitações do usuário.

Os experimentos realizados foram capazes de demonstrar que o algoritmo descrito possui um desempenho aceitável, em relação a técnicas similares, em cenários onde os nós possuem probabilidade de falhar. Ademais, o tempo de convergência do cálculo do valor médio das medições da rede e de sincronização dos nós é relativamente baixo para redes com até cinquenta nós.

Referências

- Akkaya, K. and Younis, M. (2005). A survey on routing protocols for wireless sensor networks. *Elsevier Journal of Ad Hoc Networks*, 3:325–349.
- Breza, M. and McCan, J. A. (2008). Lessons in implementing bio-inspired algorithms on wireless sensor networks. ESA Conference on Adaptive Hardware and Systems.
- Culler, D., Estrin, D., and Srivastava, M. (2004). Overview of sensors network. *IEEE Computer Magazine*, 37:41–49.
- Cunha, D. and Duarte, O. C. M. B. (2005). Um esquema bio-inspirado para estimação de campo com redes de sensores sem fio. XXII Simpósio Brasileiro de Redes de Computadores.
- Intanagonwiwat, C., Govindan, R., and Estrin, D. (2000). Directed diffusion: A scalable and robust communication paradigm for sensor networks. 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking.
- Kempe, D., Dobra, A., and Gehrke, J. (2003). Gossip-based computation of aggregate information. 44th Annual IEEE Symposium on Foundations of Computer Science.
- Lessman, J., Heimfarth, T., and Janacik, P. (2008). Shox: An easy to use simulation platform for wireless networks. 10th International Conference on Modelling & Simulation.
- Madden, S., Franklin, M. J., Hellerstein, J. M., and Hong, W. (2002). Tag: A tiny aggregation service for ad-hoc sensor networks. pages 131–146. SIGOPS Oper. Syst.
- Mills, K. (2007). A brief survey of self-organization in wireless sensor networks. *Wireless Communications and Mobile Computing*, 7:823–834.
- Nakamura, E. F., Loureiro, A. A. F., and Frery, F. G. (2007). Information fusion for wireless sensor networks: Methods, models, and classifications. *ACM Computing Surveys*, 39:1–55.
- Sumpter, D. J. T. (2006). The principles of collective animal behaviour. *Philosophical transactions of the Royal Society of London*, 361:5–22.