

Roteamento com Agregação de Dados em Redes de Sensores

Jorgito Matiuzzi Stochero, Antonio José Gonçalves Pinto e José Ferreira de Rezende

Grupo de Teleinformática e Automação – Universidade Federal do Rio de Janeiro (UFRJ) Caixa Postal 68504 21945-970– Rio de Janeiro – RJ – Brasil

{stochero, antonio, rezende}@gta.ufrj.br

***Abstract.** Sensor networks are a specialized form of mobile network where computing devices are tiny remote nodes, provided with sensing, processing, and communication capabilities. These devices collect data from the environment, interact with each other and pass it along until it reaches the users. They differ from traditional networks because sensor nodes are densely deployed, prone to failures and task specific, imposing new challenges for routing protocols. These networks are data centric, because data is more important to the network than the nodes themselves, and the protocols should take this into account during routing and addressing. In this paper we describe sensor network routing protocols and propose a technique based on bayesian inference to aggregate multisensor data while routing, reducing message traffic in order to save energy and extend the lifetime of the network.*

***Resumo.** Redes de sensores são uma forma especializada de redes móveis onde os dispositivos computacionais são pequenos nós remotos, capazes de sensoriar, processar e comunicar dados. Estes dispositivos coletam dados do ambiente, interagem uns com os outros até que o dado chegue ao usuário. Tais redes diferem das tradicionais porque seus nós são densamente distribuídos, sujeitos a falhas e atuam para tarefas específicas, o que impõe desafios aos protocolos de roteamento. Estas redes são centradas em dados, uma vez que a informação é mais importante que o nó que a obtém, o que deve ser levado em conta durante o roteamento e o endereçamento. Neste artigo descrevemos protocolos de roteamento para redes de sensores e propomos uma técnica baseada em inferência bayesiana para agregar dados de múltiplos sensores durante o roteamento, reduzindo o tráfego de mensagens para economizar energia e aumentar o tempo de vida da rede.*

1. Introdução

Recentes avanços em micro-eletrônica e em comunicações sem fio possibilitaram a implementação de um novo tipo de redes de computadores móveis: redes de sensores. É previsto que, em um futuro próximo, dispositivos sensores, dotados de componentes de sensoriamento, processamento e comunicações de dados em um único circuito integrado estarão disponíveis no mercado a um baixo custo. Estes dispositivos deverão ser menores do que 1cm³, pesar menos do que 100 g, e devem consumir pouca energia, para evitar troca de baterias [1].

Uma rede de sensores é composta de centenas a milhares desses dispositivos (*nós*) dispostos em uma área geográfica de interesse. Cada nó coleta dados do ambiente, transforma a informação obtida em uma descrição do fenômeno observado, e gerencia os protocolos de comunicação. Devido ao seu pequeno tamanho, os nós são colocados próximos ao fenômeno a ser monitorado, aumentando a acurácia da observação. Suas aplicações podem ser militares, na detecção e monitoração de movimentos inimigos em uma área inóspita, médicas, na monitoração remota de parâmetros vitais de um paciente, entre outras.

Os protocolos devem possuir características de auto-configuração para descobrir quais os nós possuem informação (fontes), quais precisam dessa informação (destinos ou *sorvedouros*), e qual a melhor forma de transferir a informação de um para o outro, com mínimo consumo de energia. De acordo com o modelo rádio de primeira ordem proposto em [2], comunicações de dados são as responsáveis pelo maior peso no consumo de energia da rede, em comparação com o sensoriamento e o processamento. Dessa forma, para economizar energia é interessante aumentar o processamento dos dados para diminuir o número de transmissões, bem como utilizar transmissões por distâncias curtas (dezenas de metros) entre os nós, obtendo com isso uma menor potência no transmissor. Como em geral os sorvedouros estão bem afastados das fontes, isto significa que um pacote de dados será transmitido ao longo de vários saltos usando nós intermediários. Este relacionamento muitos-para-um (fontes para sorvedouros), bem como o fato de que normalmente as fontes observam o mesmo fenômeno, sugere redundância nos dados sendo transmitidos.

O objetivo deste artigo é explorar esta redundância utilizando, durante o roteamento, uma técnica de agregação para minimizar a transmissão de dados na rede e com isso aumentar o seu tempo de vida. O artigo é organizado da seguinte forma: na seção 2 são discutidos algoritmos de roteamento em redes de sensores; na seção 3 são apresentadas técnicas de agregação de dados; a seção 4 traz a proposta do trabalho, a simulação realizada e os resultados obtidos e o trabalho é finalizado na seção 5.

2. Roteamento em Redes de Sensores

A figura 1 mostra uma rede de sensores típica.

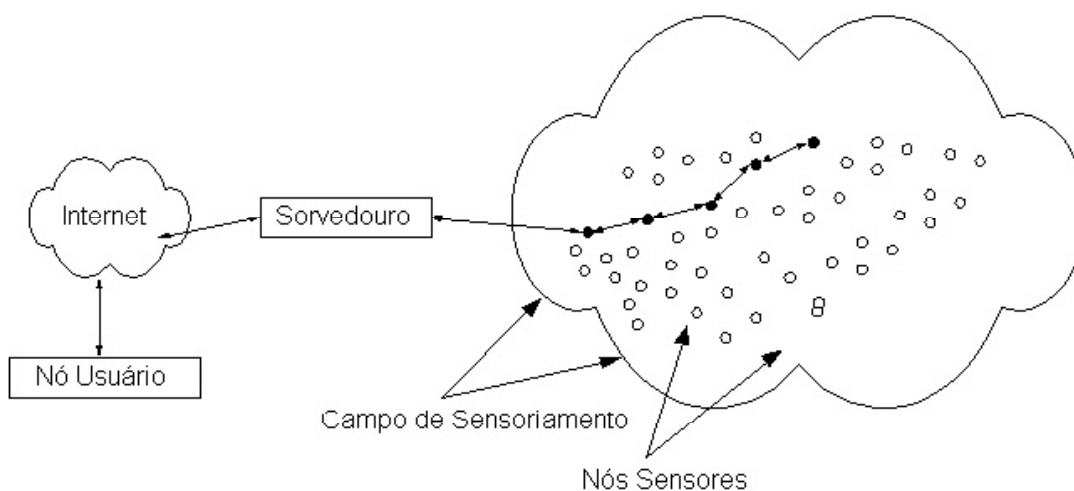


Figura 1. Rede de Sensores

O sorvedouro pode ser o consumidor final dos dados coletados, ou apenas um *gateway* para uma rede externa (Internet). Em geral, a utilização da rede de sensores depende da aplicação. O sorvedouro recebe as tarefas de sensoriamento da aplicação e as difunde até que as fontes iniciem as suas atividades. Sendo assim, o fluxo de tarefas vai do sorvedouro até as fontes, enquanto os dados fluem no sentido inverso. Dependendo do estado da rede, o mesmo nó pode atuar ora como sorvedouro, ora como fonte, caracterizando uma topologia de rede flexível e específica para cada aplicação.

Existem duas famílias de protocolos de roteamento aplicáveis para redes de sensores: Hierárquicos ou planos. A figura 2 ilustra as diferenças:

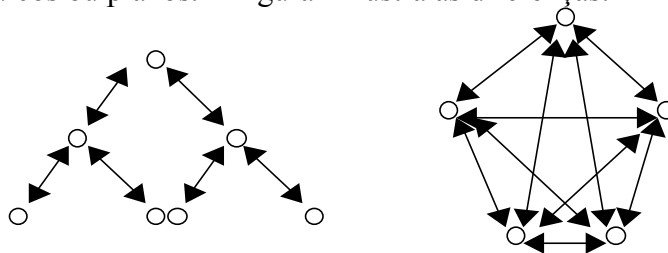


Figura 2. Protocolos de roteamento planos e hierárquicos

Na figura da esquerda é observada a topologia hierárquica, caracterizada pela existência de grupamentos de nós (*clusters*). Cada nó dentro do *cluster* comunica-se apenas com o seu pai (*cluster head*), que pode formar *clusters* de ordem mais alta. Cada *cluster head* integra a informação dos nós abaixo dele por meio de alguma técnica de agregação e reporta os resultados acima na hierarquia. O nó de mais alto nível é chamado de estação base e é responsável pela entrega da informação ao usuário. Os problemas nesta classe de protocolos estão na seleção dos *cluster heads* e na configuração da hierarquia de forma a garantir um balanceamento no gasto de energia dos nós. Exemplos de protocolos hierárquicos incluem o LEACH [2] e o PEGASIS[3].

Na figura da direita é observada a topologia plana. Cada nó é autônomo e conectado a alguns ou a todos os outros nós, dependendo do alcance do seu rádio. Não existem *cluster heads* ou estações bases. Esta abordagem causa um pesada carga nas comunicações da rede, pois são necessárias n^2 conexões para uma rede de n nós. Para evitar isso, os protocolos devem identificar e privilegiar os melhores caminhos das fontes ao destino e limitar dessa forma o número de conexões. Exemplos de protocolos planos são SPIN [4], EAR [5], e Difusão Direcionada, ou *Directed Diffusion* [6].

O roteamento deve considerar ainda dois outros aspectos: endereçamento e o fato de redes de sensores serem centradas em dados, ou seja, o dado a ser transmitido é mais importante do que o nó que o transmite. O elevado número de nós e os escassos recursos de processamento tornam a utilização de um identificador (ID) global único para cada nó um aumento nos custos de transmissão. Sendo assim, os nós podem ser reconhecidos pelo tipo de informação que provêm (como temperatura, presença de fumaça, etc), caracterizando o que se chama de endereçamento baseado em atributos, facilitando a agregação de dados perto de onde eles são gerados. Exemplos de esquemas de endereçamento baseado em atributos podem ser achados em [7] e [8].

Um modelo para roteamento centrado em dados foi sugerido em [9], onde foi comparado com o centrado em endereços. Em [9], os nós intermediários lêem o conteúdo da mensagem e processam alguma função de agregação de pacotes, reduzindo o número de transmissões. No roteamento centrado em endereços, cada fonte envia os

dados de forma independente ao destino. A agregação pode adicionar retardo, porque os dados de fontes próximas precisam esperar por dados de fontes mais distantes para serem combinados. Algumas técnicas de agregação são discutidas a seguir.

3. Técnicas de Agregação de Dados

Na agregação de dados, uma rede de sensores é vista como uma árvore reversa, na qual o sorvedouro pede aos sensores que relatem as condições do fenômeno observado. Os dados que chegam de múltiplos sensores são agregados a cada nó com base nos atributos observados. Em [4] a agregação é explicada como uma técnica usada para solucionar o problema da implosão (*implosion*) e da sobreposição (*overlap*) em roteamento centrado nos dados.

As técnicas utilizadas para combinar ou agregar dados podem ser triviais como as funções exploradas em [10] (“máximo”, “média”, “contador”, etc.), adequadas para fusão de dados de um mesmo tipo de sensor, ou complexas como as apresentadas em [11], que realizam a fusão de dados de múltiplos sensores. A maioria das aplicações de agregação discutidas em pesquisas de redes de sensores são baseadas em funções simples. Nesse trabalho estudamos um cenário mais complexo, com o uso de técnicas de fusão de dados de diferentes tipos de sensores, descritas a seguir: inferência clássica, inferência bayesiana, os métodos Dempster-Shafer e heurísticas.

A inferência clássica procura determinar a validade das hipóteses propostas (versus hipóteses alternativas) baseada em probabilidades empíricas. Calcula a probabilidade conjunta dada uma hipótese assumida. Em geral pode validar somente duas hipóteses de cada vez (hipótese nula e sua alternativa) e não considera a informação a priori da crença da hipótese proposta.

A inferência bayesiana atualiza as probabilidades das hipóteses alternativas baseada nas observações de evidências. A fórmula de Bayes fornece uma relação entre a probabilidade a priori, a probabilidade condicional de uma observação dada uma hipótese e a probabilidade a posteriori de uma hipótese. Novas informações são utilizadas para atualizar a probabilidade a priori da hipótese em um procedimento análogo a um experimento científico. Probabilidades empíricas ou subjetivas podem ser utilizadas e várias hipóteses podem ser avaliadas simultaneamente.

O método Dempster-Shafer (DS) é uma generalização da inferência bayesiana. O método DS relaxa a restrição bayesiana quanto as hipóteses mutuamente exclusivas, adicionando um grau de incerteza.

As heurísticas utilizam várias técnicas não probabilísticas. Alguns métodos imitam a forma como humanos realizam inferência. São exemplos dessas técnicas os algoritmos de votação e técnicas de consenso, entre outras.

4. Proposta, Simulação e Resultados

Utilizamos o protocolo de roteamento *Directed Diffusion* [6], com a adição de um módulo para agregação de dados na rede utilizando inferência bayesiana. A rede é constituída por dois tipos de nós fonte, um para medir a temperatura, e outro para detectar fumaça. A aplicação necessita detectar fogo na área e a rede procura agregar as informações que colaborem com a hipótese de fogo antes de que cada mensagem (fumaça ou temperatura) chegue individualmente ao sorvedouro. Nesta seção nós avaliamos o desempenho dessa rede e verificamos se existe algum ganho com a proposta.

4.1 Fundamentação Teórica

A inferência bayesiana é baseada no teorema de Bayes. Supondo que H_1, H_2, \dots, H_j representam hipóteses mutuamente exclusivas e exaustivas que podem explicar um evento E (uma observação) que acabou de ocorrer, temos:

$$\text{onde, } p(H_i/E) = \frac{p(E/H_i)p(H_i)}{\sum_i p(E/H_i)p(H_i)} \qquad \sum_i p(H_i) = 1$$

$p(H_i/E)$ = probabilidade a posteriori da hipótese H_i ser verdadeira dado a evidência E

$p(H_i)$ = probabilidade a priori da hipótese H_i ser verdadeira

$p(E/H_i)$ = probabilidade de observar a evidência E , dado que H_i é verdadeira

4.2 Cenário de Simulação

O cenário de simulação consiste em verificar a hipótese *fogo* levando-se em consideração as evidências *temperatura* e *fumaça*. Assumindo serem conhecidas as probabilidades a priori das três variáveis, a fórmula de Bayes é usada para calcular, a partir das evidências coletadas no local, a probabilidade de existir fogo. Para minimizar o impacto do retardo da rede, a agregação é realizada nó a nó com a informação disponível, ou seja, o nó utiliza um cache local para fazer a agregação, caso já tenha recebido uma mensagem similar. Ele não aguarda a chegada de novas mensagens. Foi utilizado o simulador ns-2 [12] e o campo de sensoriamento consistiu de uma área de 670 x 670 m, onde foram dispostos aleatoriamente 100 nós sensores. Cada nó possui um raio de cobertura de 250 m e as fontes e o sorvedouro são posicionados afastados.

4.3 Resultados

Embora a formação de uma árvore de agregação ótima é geralmente um problema NP-completo [9], existem árvores geradas heurísticamente que possibilitam prolongar o tempo de vida da rede. As simulações iniciais apresentaram significativa redução no número de comunicações, e portanto, economia de energia, como pode ser visto nos gráficos da figura 3 para simulações de 500 e 1000 segundos.

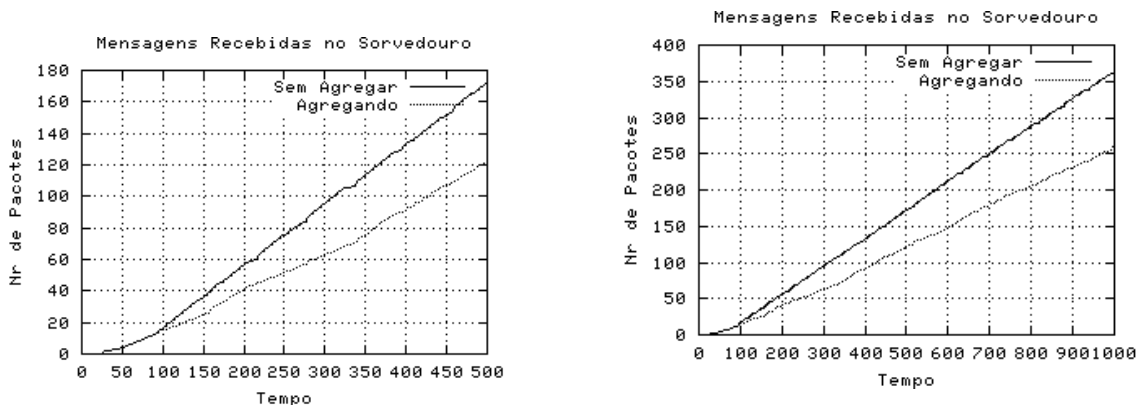


Figura 3. Topologia de 100 nós, campo de sensoriamento de 670 x 670 m

5. Conclusão e Trabalhos Futuros

Neste artigo descrevemos as características principais das redes de sensores, especialmente no que diz respeito ao roteamento de dados. Como o principal fator limitante ao tempo de vida útil de uma rede de sensores é a durabilidade das baterias dos seus nós, os protocolos de roteamento devem buscar a minimização do tráfego de mensagens e, para isso, aproveitar-se de endereçamento baseado em atributos e nas características centradas em dados da rede. Utilizamos o protocolo *Directed Diffusion* com o acréscimo de um módulo para agregação de dados utilizando inferência bayesiana e nas simulações efetuadas observamos a diminuição do tráfego de mensagens das fontes ao destino.

Como próximos trabalhos, pretendemos avaliar o desempenho de outras técnicas de agregação e compará-las com os resultados obtidos com a inferência bayesiana. Pretendemos também acrescentar um temporizador em cada nó para aguardar novas mensagens antes de agregar, aumentando o descarte de pacotes, verificando a influência no retardo sofrido pelos dados desde a fonte até o sorvedouro. Outro foco de interesse é em modificar o estabelecimento de rotas no *Directed Diffusion* para beneficiar nós intermediários com potencial para agregação e com isso aumentar os ganhos em energia com uma redução mais acentuada do tráfego.

6. Referências

- [1] J. Rabaey, J. Ammer, J. L. da Silva Jr., D. Patel. "PicoRadio: Ad-hoc Wireless Networking of Ubiquitous Low-Energy Sensor/Monitor Nodes". IEEE WVLSI 2000.
- [2] W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks" Proceedings of the 33rd Hawaii International Conference on System Sciences, January 2000.
- [3] S. Lindsey, C. S. Raghavendra, "PEGASIS: Power Efficient Gathering in Sensor Information Systems," IEEE ICC 2001.
- [4] Joana Kulik, Wendi Heinzelman, and Hari Balakrishnan. "Negotiation-based Protocols for Disseminating Information in Wireless Sensor Networks". IEEE/ACM Mobicom 1999.
- [5] Rahul C. Shah and Jan Rabaey, "Energy Aware Routing for Low Energy Ad Hoc Sensor Networks", IEEE Wireless Communications and Networking Conference 2002.
- [6] Chalermek Intanagonwiwat, Ramesh Govindan, and Deborah Estrin. "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks". Mobicom, 2000.
- [7] J. Heidemann, F. Silva, C. Intanagonwiwat, R. Govindan, D. Estrin, and D. Ganesan, "Building Efficient Wireless Sensor Networks with Low-Level Naming". 18Th ACM Symposium on Operating Systems Principles, 2001.
- [8] William Adjie-Winoto, Elliot Schwartz, Hari Bvalakrishnan, and Jeremy Lilley. "The design and implementation of an intentional naming system". 17Th ACM Symposium on Operating Systems Principles, 1999.
- [9] Bhaskar Krishnamachari, Debora Estrin, Stephen Wicker. "Modeling Data-Centric Routing in Wireless Sensor Networks". IEEE INFOCOM 2002.
- [10] A. Boulis, S. Ganeriwal, M. Srivastava. "Aggregation in Sensor Networks: An Energy-Accuracy Trade-off". Sensor Network Protocols and Applications, Special Issue of Elsevier Ad Hoc Networks Journal, 2003.
- [11] L. A. Klein. "Sensor and Data Fusion Concepts and Applications". Second Edition. Tutorial Texts in Optical Engineering SPIE PRESS Vol. TT355, 1999
- [12] The Network Simulator ns-2, <http://www.isi.edu/nsnam/ns>, acessado em 21 Ago 2003.