MECANISMO DE AGREGAÇÃO DE DADOS EMPREGANDO TÉCNICAS PARAMÉTRICAS EM REDES DE SENSORES

Antonio José Gonçalves Pinto

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA ELÉTRICA.

Aprovada por:

Prof. José Ferreira de Rezende, Dr.

Prof. Julius Cesar Barreto Leite, Ph. D.

Prof. Otto Carlos Muniz Bandeira Duarte, Dr. Ing.

RIO DE JANEIRO, RJ - BRASIL

JUNHO DE 2004

PINTO, ANTONIO JOSÉ GONÇALVES

Mecanismo de Agregação de Dados Empregando Técnicas Paramétricas em Redes de Sensores [Rio de Janeiro] 2004

XVI, 99 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia Elétrica, 2004)

Tese - Universidade Federal do Rio de Janeiro, COPPE

1. Redes de Sensores
2. Agregação de Dados
3. Inferência Bayesiana
4. Inferência de Dempster-Shafer

I. COPPE/UFRJ II. Título (série)

À minha esposa Claudia e ao meu filho Guilherme, pelo amor, compreensão e carinho dispensados enquanto me dediquei a este trabalho.

Agradecimentos

Ao Altíssimo, por guiar os meus passos em todos os momentos da minha vida.

Ao professor José Ferreira Rezende, não só pela valiosa orientação dos trabalhos, mas também, pela dedicação, disponibilidade e exemplo evidenciados durante todo o meu trabalho de pesquisa.

À professora Marly (UFAM) e ao professor Eduardo (UFRJ), pela inciativa e apoio para realização do curso de mestrado.

A todos os companheiros do Sistema de Vigilância da Amazônia (SIVAM), pelo apoio e disposição em ajudar.

Ao Centro de Desenvolvimento de Sistemas (CDS) do Exército Brasileiro, pelo tempo disponibilizado para o desenvolvimento deste trabalho e estímulo ao aperfeiçoamento de seu corpo técnico.

Ao amigo Stochero, que desenvolveu sua pesquisa paralelamente à minha e com quem discuti muitas idéias.

Aos companheiros do Grupo de Teleinformática e Automação (GTA)/COPPE da Universidade Federal do Rio de Janeiro e do Núcleo de Tecnologia Eletrônica e da Informação (NUTELI) da Universidade Federal do Amazonas, pelos momentos compartilhados.

Ao CNPq, CAPES, COFECUB, FAPERJ e SUFRAMA, pelos recursos utilizados para essa pesquisa.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

MECANISMO DE AGREGAÇÃO DE DADOS EMPREGANDO TÉCNICAS PARAMÉTRICAS EM REDES DE SENSORES

Antonio José Gonçalves Pinto

Junho/2004

Orientador: José Ferreira Rezende

Programa: Engenharia Elétrica

A organização de dispositivos microsensores em uma sofisticada arquitetura computacional e de comunicação permitiu o desenvolvimento de um novo tipo de rede móvel sem fio, denominada rede de sensores. Este tipo especial de redes *Ad Hoc* atende a requisitos específicos e tem a conservação de energia como um dos fatores mais importantes a serem considerados no projeto da rede. Nesse aspecto, a agregação de dados é fundamental para prolongar o tempo de vida da rede, realizando um processamento eficiente enquanto os dados fluem das fontes para o sorvedouro. Este trabalho propõe um mecanismo de agregação de dados baseado em técnicas paramétricas, de forma a proporcionar uma economia de energia, ao mesmo tempo em que dados brutos são transformados em informações de interesse, com um baixo custo na rede. O mecanismo explora a característica de protocolos de roteamento centrados em dados na formação e manutenção da árvore de agregação, para aplicar as técnicas da Inferência Bayesiana e de Dempster-Shafer, que permitem um mapeamento direto entre as características do fenômeno com o evento que se quer observar. O desempenho do mecanismo é avaliado por meio de simulações e os resultados mostraram uma significativa economia de energia.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

DATA AGGREGATION MECHANISM USING PARAMETRIC TECHNICS IN SENSOR NETWORKS

Antonio José Gonçalves Pinto

June/2004

Advisor: José Ferreira Rezende

Department: Electrical Engineering

The organization of microsensors devices in computational and communication sophisticated architecture allows the development of a new type of wireless network, named sensor networks. This special type of *Ad Hoc* networks have specifics requirements and the matter of conservation of energy is one of the most important factors to be considered in a design of these networks. In this context, data aggregation is fundamental to prolonging the lifetime of the network by doing an efficient process while data flow from sources to sink. This work proposes one aggregation mechanism based on parametric techniques in order to save energy at same time that raw data are transformed in information of interest, with a low cost to the network. The mechanism explores the feature of data centric route protocols in the aggregation tree building and maintenance to apply Bayesian Inference and Dempster-Shafer techniques that allow directly map parametric data into an event of interest. The performance of mechanism is evaluated by means of simulations and the results showed a significant saving energy.

Lista de Acrônimos

AODV: *Ad Hoc On-Demand Distance Vector Routing*;

API : *Application Program Interface*;

ARQ: *Automatic Repeat Request*;

ASIC: *Application Specific Integrated Circuit*;

BPSK: *Binary Phase Shift Keying*;

CDMA: *Code Division Multiple Access*;

CSMA: *Carrier Sense Multiple Access*;

DSDV: *Destination-Sequenced Distance-Vector Routing*;

DSR: *Dynamic Source Routing*;

ESRT: *Event-to-Sink Reliable Transport*;

FEC: *Forward Error Correction*;

FPGA: *Field Programmable Gate Array*;

GPS: *Global Positioning System*;

IEEE: *Institute of Electrical and Electronics Engineers*;

IETF: *Internet Engineering Task Force*;

IP: *Internet Protocol*;

ISM: *Industrial, Scientific, and Medical*;

ITU-T: *International Telecommunication Union - Telecommunications Standardization Sector*;

JDL DFS: *Joint Directors of Laboratories Data Fusion Subpanel*;

LEACH: *Low-Energy Adaptive Clustering Hierarchy*;

MAC: *Medium Access Control*;

MANET: *Mobile Ad Hoc Network*;

MBWA: *Mobile Broadband Wireless Access*;

MECN: *Minimum Energy Communication Network*;

RMST: *Reliable Multi-Segment Transport*;

RSSF: Rede de Sensores sem Fio;

SAR: *Sequential Assignment Routing*;

SINA: *Sensor Information Networking Architecture and Applications*;

SMACS: *Self-Organizing Medium Access Control for Sensor Networks*;

SNEP: *Secure Network Encryption Protocol*;

SPIN: *Sensor Protocols for Information via Negotiation*;

SPINS: *Security Protocols for Sensor Networks*;

SQTL: *Sensor Query and Tasking Language*;

TCP: *Transmission Control Protocol*;

TDMA: *Time Division Multiple Access*;

TESLA: *Timed, Efficient, Streaming, Loss-tolerant Authentication Protocol*;

UDP: *User Datagram Protocol*;

USC/ISI: *University of Southern California/Information Sciences Institute*;

UWB: *Ultra Wideband*;

WINS: *Wireless Integrated Networks Sensors*;

WirelessMAN: *Broadband Wireless Metropolitan Area Network*;

WLAN: *Wireless Local Area Network*;

WPAN: *Wireless Personal Area Network*;

WRP: *Wireless Routing Protocol.*

Sumário

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

1.1 Motivação

A crescente evolução das técnicas de comunicação, especialmente na área de processamento de sinais e de microeletrônica permitiu o desenvolvimento de pequenos dispositivos com baixo consumo de energia capazes de combinar as capacidades de sensoriamento e de comunicação sem fio. A organização de tais dispositivos microsensores em uma sofisticada arquitetura computacional e de comunicação permitiu criar um novo tipo de rede sem fio, denominada rede de sensores sem fio (RSSF). Essa nova tecnologia permite monitorar remotamente um ambiente e obter informações mais precisas, tirando vantagem de ser empregada bem próxima ao fenômeno de interesse.

A diversidade e o potencial vislumbrado das aplicações que empregam essa nova tecnologia têm motivado consideravelmente a pesquisa na área. As aplicações de RSSFs estendem-se por diversas áreas de interesse como, por exemplo, na medicina, no controle ambiental, em ações de resgate, no rastreamento de alvos militares, na segurança de instalações, dentre outras possíveis.

Uma das características mais importantes em uma RSSF é o consumo de energia. O tempo de vida da rede depende da eficiência em termos de energia de cada solução de *hardware*, *software* e *middleware* que compõe sua arquitetura. A comunicação entre os nós é a maior responsável pelo consumo de energia quando comparada com o sensoriamento e processamento. Por esta razão, esta comunicação deve ocorrer a curtas distâncias, obrigando que os pacotes de dados sejam encaminhados por meio de rotas com múltiplos saltos. Esse tipo de roteamento sugere o aumento do processamento em cada nó sensor para minimizar as transmissões. O processamento distribuído na rede com o objetivo de reduzir a quantidade de dados é totalmente explorado pela agregação de dados.

A agregação consiste em combinar dados oriundos de diferentes nós sensores para ampliar o entendimento do ambiente monitorado e, ao mesmo tempo, eliminar as transmissões redundantes. O sensoriamento de um fenômeno comum e a alta densidade dos nós, características comuns das RSSFs, podem gerar um elevado nível de redundância, o que favorece a agregação de dados durante o roteamento destes.

A agregação de dados em RSSFs envolve três diferentes aspectos: *onde*, *como* e *quando* agregar [2]. O primeiro aspecto, *onde* agregar, é determinado pelo protocolo de roteamento, enquanto os outros dois são resolvidos pelo mecanismo de agregação, que define qual técnica será empregada e o momento em que os dados serão combinados. O protocolo de roteamento nas RSSFs é responsável pela construção de uma árvore que interliga vários nós fontes a um nó sorvedouro (árvore $N \mapsto 1$), na qual os nós comuns dos diversos caminhos independentes constituem os pontos de agregação. O mecanismo de agregação pode valer-se de uma função trivial que realiza a supressão de mensagens redundantes ou outros métodos básicos, que incluem calcular o valor máximo, mínimo ou médio do conteúdo de mensagens com a mesma semântica. Existem, ainda, as técnicas paramétricas oriundas do modelo de fusão de dados criado pelo *Joint Directors of Laboratories Data Fusion Subpanel* (JDL DFS), que permitem combinar dados gerados por sensores de diferentes tipos. O modelo JDL DFS foi desenvolvido para fusão de dados de sensores tradicionais como, por exemplo, radares e imageadores espectrais, não sendo diretamente aplicado em RSSFs.

Diversos pesquisadores têm explorado os impactos da agregação de dados em RSSFs [3, 4, 5, 6]. Os resultados obtidos mostram que uma significante economia de energia pode ser obtida com a agregação. Entretanto, a necessidade de coordenação entre os nós e o retardo introduzido na rede pelos procedimentos tradicionais de armazenar e reencaminhar dados não devem ser negligenciados, sob o risco de comprometer o desempenho da rede.

As RSSFs constituem uma aplicação recente para as redes de comunicação sem fio *Ad Hoc*. Portanto, as arquiteturas atuais propostas para as redes *Ad Hoc* não atendem completamente as necessidades das RSSFs, carecendo de soluções específicas para os problemas que envolvem a dinâmica de operação da rede como, por exemplo, mecanismos eficientes de agregação e compartilhamento de tarefas entre nós vizinhos. Intanagonwiwat *et al.* [7] fazem uma analogia situando o estado da arte atual das RSSFs com o da Internet de três décadas atrás.

Desta forma, a principal motivação para o estudo de um mecanismo de agregação envolvendo técnicas paramétricas deve-se ao fato do seu grande potencial de emprego destacado na literatura de RSSFs. Embora sua eficiência tenha sido comprovada em diversos estudos com sensores tradicionais, sua aplicação em RSSFs ainda necessita de pesquisa e avaliação de desempenho.

1.2 Objetivos

Esta dissertação concentra-se no mecanismo de agregação de dados empregando técnicas paramétricas em RSSFs. Tal mecanismo explora a formação da árvore de agregação em um protocolo de roteamento centrado em dados. O protocolo de roteamento escolhido para implementação e avaliação de desempenho deste mecanismo foi a Difusão Direcionada (*Directed Diffusion*) [8]. Na RSSF baseada na Difusão Direcionada, os nós trabalham de forma cooperativa para armazenar e processar as mensagens. Este protocolo caracteriza-se pela robustez, capacidade de coordenação de um elevado número de nós e independência da topologia da rede. Outro importante aspecto desse paradigma de comunicação diz respeito às interações localizadas (isto é, baseadas apenas em informações locais trocadas entre nós vizinhos) que são responsáveis por todas as decisões do nó sensor. A manutenção dessas características foi um requisito básico do mecanismo de agregação de dados desenvolvido neste trabalho.

As técnicas paramétricas de fusão de dados permitem o mapeamento direto entre as características do fenômeno e o evento que se quer observar. Por exemplo, seja o mapeamento simplificado entre o evento *fogo* e aquilo que o particulariza: *temperatura*, *umidade*, *fumaça*, *monóxido de carbono* e *dióxido de carbono*. Esse mapeamento tira proveito do esquema de endereçamento baseado em atributos, utilizado pela Difusão Direcionada, ao difundir na rede os interesses pelas características do fenômeno. As técnicas paramétricas estudadas incluem a Inferência Bayesiana e de Dempster-Shafer, que empregam métodos estatísticos e heurísticos para integrar dados.

Os objetivos desta dissertação são o estudo do paradigma da Difusão Direcionada e das técnicas paramétricas de fusão de dados, visando a adequação de um mecanismo de agregação de dados utilizando a Inferência Bayesiana e de Dempster-Shafer em RSSFs. O mecanismo de agregação de dados tem por fim proporcionar economia de energia por meio da redução da comunicação entre os nós, ao mesmo tempo em que dados brutos são transformados em informações de interesse, minimizando os impactos na latência e robustez da rede.

Para avaliar a adequação e o desempenho do mecanismo de agregação estudado foram desenvolvidos dois filtros, chamados Filtros Bayes e DS, que foram adicionados ao módulo da Difusão Direcionada disponível no simulador *ns-2* [9]. Estes módulos estão disponíveis na seguinte URL: http://www.gta.ufrj.br/wsns.

As análises iniciais do mecanismo de agregação proposto, realizadas pelo autor deste trabalho de dissertação, foram publicadas em [10]. Uma análise com mais detalhes desse mecanismo, também será publicada em [11].

As principais contribuições deste trabalho consistem na proposta e avaliação de desempenho de um mecanismo de agregação que difere do procedimento tradicional de armazenar e reencaminhar mensagens. Neste mecanismo foram aplicadas e adaptadas as técnicas de Inferência Bayesiana e de Dempster-Shafer para agregação de diferentes tipos de dados em RSSFs.

1.3 Organização da Dissertação

A dissertação está organizada em seis capítulos. Os primeiros três capítulos fazem a apresentação do tema de pesquisa e fornecem o embasamento teórico do trabalho. Nos capítulos seguintes são abordados a proposta do mecanismo de agregação empregando técnicas paramétricas em RSSFs e os resultados obtidos com as simulações. A organização detalhada é descrita a seguir:

- Capítulo 1 - Faz a introdução da dissertação, apresentando o tema de pesquisa;
- Capítulo 2 - Aborda os conceitos básicos relacionados à RSSF, importantes para o entendimento desta dissertação;
- Capítulo 3 - Apresenta as questões diretamente relacionadas aos objetivos da pesquisa, envolvendo o protocolo de roteamento da Difusão Direcionada e as técnicas paramétricas de fusão de dados (a Inferência Bayesiana e de Dempster-Shafer);
- Capítulo 4 - Descreve o mecanismo proposto na dissertação, explicando os aspectos *onde*, *como* e *quando* agregar dados em uma RSSF;
- Capítulo 5 - Apresenta o ambiente de simulação e os cenários escolhidos para avaliar os Filtros Bayes e DS, assim como, os principais resultados obtidos; e
- Capítulo 6 - Traz as conclusões da dissertação e sugestões de continuação do trabalho.

Capítulo 2

Conceitos Básicos

A RSSF é uma evolução das redes de comunicação sem fio *Ad Hoc*, ou MANETs (*Mobile Ad Hoc Networks*) [12], que não necessitam de qualquer infra-estrutura pré-existente para prover comunicação entre os nós. No entanto, a arquitetura atual das redes sem fio *Ad Hoc* não atende completamente às necessidades das RSSFs, considerando que estas possuem um maior número de nós, uma elevada taxa de falhas e são extremamente limitadas em consumo de energia. As RSSFs possuem requisitos únicos que demandam uma arquitetura de rede específica, cuja dinâmica de operação estabelece enormes desafios para os protocolos de comunicação.

Este capítulo aborda conceitos básicos relacionados às RSSFs, importantes para o correto entendimento desta dissertação. A seção 2.1 traz considerações sobre redes sem fio em geral. A seção 2.2 descreve seus principais requisitos e características. A seção 2.3 fornece uma visão sobre a arquitetura dessas redes. A seção 2.4 aborda a pilha de protocolos para a completa realização das tarefas de comunicação e sensoriamento. Finalmente, a seção 2.5 conclui o capítulo.

2.1 Redes Móveis Sem Fio

Uma rede móvel sem fio consiste em uma coleção de terminais com transmissores e receptores que se movimentam em uma determinada área, geralmente utilizando transmissões com freqüência de rádio ou infravermelho [13]. As tecnologias de comunicação de dados sem fio disponíveis atualmente foram desenvolvidas para atender requisitos diversificados, cada uma delas constituindo um padrão para comunicação sem fio. Esses padrões são criados por órgãos de padronização internacional da área de telecomunicações como, por exemplo, o ITU-T (*International Telecommunication Union - Telecommunications Standardization Sector*) [14] e o IEEE (*Institute of Electrical and Electronics Engineers*) [15].

O IEEE constitui a maior organização de padronização do mundo [13], possuindo quatro grupos de trabalhos (*Working Groups*) dedicados a redes sem fio. O Grupo de Trabalho 11 [16] é responsável pelo padrão 802.11, para redes locais sem fio (*Wireless Local Area Networks - WLANS*) e suas extensões 802.11a [17] e 802.11b [18]. O Grupo de Trabalho 15 [19] responde pelas redes de área pessoal (*Wireless Personal Area Networks - WPANs*). O Grupo de Trabalho 16 [20] elabora as especificações para redes metropolitanas sem fio (*Broadband Wireless Metropolitan Area Networks - WirelessMANs*). E finalmente, o Grupo de Trabalho 20 [21] trabalha para padronizar o tráfego nativo IP para acesso em banda larga de forma móvel (*Mobile Broadband Wireless Access - MBWA*).

De acordo com o padrão 802.11, as redes móveis sem fio podem ser classificadas como infra-estruturadas e *Ad Hoc*. Na rede sem fio infra-estruturada, a área de cobertura é dividida em regiões menores, denominadas células. Em cada região é necessário alocar ao menos um ponto de acesso, também denominado de estação rádio base, para fornecer o serviço de rede para os nós móveis. O nó móvel conecta-se à rede para estabelecer uma conexão sem fio com a estação base, enquanto estas comunicam-se entre si através de uma rede fixa de alta velocidade. Na rede *Ad Hoc*, todos os nós movem-se livremente e não necessitam de nenhuma infra-estrutura de comunicação. Os nós são conectados diretamente entre eles, atuando como roteadores em uma topologia que tende a mudar com freqüência. Cada nó utiliza um protocolo de roteamento que permite descobrir rotas para qualquer estação da rede. O principal objetivo de uma rede *Ad Hoc* é formar e manter conectados os nós para transportar dados através de múltiplos saltos [22].

A arquitetura das redes *Ad Hoc* foi definida pelo grupo de trabalho sobre MANETs na IETF [12], cujas principais vantagens são o rápido emprego, a alta conectividade, a elevada tolerância a falhas e a mobilidade. No entanto, as redes *Ad Hoc* estudadas até o presente, consideram uma magnitude da ordem de dezenas ou centenas de nós móveis de tamanho moderado, onde a restrição de energia não é tão impositiva quanto em RSSFs. Este cenário é modificado com a introdução de uma nova aplicação que emprega nós, denominados microsensores, caracterizados pela limitada capacidade de comunicação, de processamento, de armazenamento e de consumo de energia [23]. Estes microsensores são conectados através do meio sem fio para formar um novo tipo de rede *Ad Hoc*, que atende a requisitos específicos e tem seu tempo de vida dependente de um judicioso consumo de energia [24]. Este tipo especial de rede *Ad Hoc* é denominado de rede de sensores sem fio (RSSF), objeto de estudo desta dissertação.

2.2 Redes de Sensores

Uma RSSF é um sistema distribuído que consiste de um elevado número de nós sensores (tais como, termômetro, manômetro, sísmico, detector de luz, químico e biológico) que estão conectados pelo meio sem fio [25]. A principal vantagem de uma RSSF é a capacidade de ser empregada, sem a assistência do homem, em quase todos os tipos de ambiente, sendo que a posição exata dos nós não é predeterminada, ou seja, ocorre de maneira aleatória. Por exemplo, uma RSSF poderia ser empregada no interior de máquinas, no fundo do oceano, em edificações, em um campo contaminado (química ou biologicamente), ou ainda, em um campo de batalha.

O emprego das RSSFs possui três aspectos fundamentais: o sensoriamento distribuído, a comunicação sem fio e o processamento distribuído [26]. Quando a localização precisa da região de interesse é desconhecida, o sensoriamento distribuído permite posicionar os sensores mais próximos do fenômeno do que seria possível com um único sensor, aumentando assim a resolução da região monitorada. As aplicações de sensoriamento em áreas sem infra-estrutura de comunicação, ou de energia, requerem o emprego de comunicação sem fio e fontes de energia autônomas. Finalmente, as severas restrições de energia impostas, favorecem processamento distribuído na rede, de forma a reduzir a comunicação entre os nós.

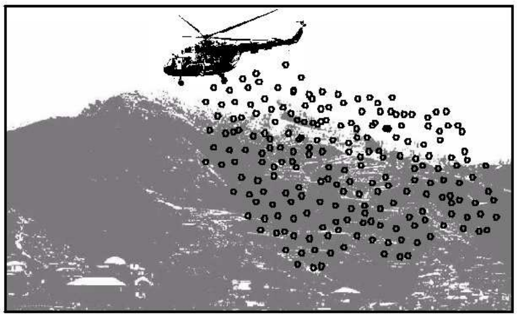
As aplicações das RSSFs abrangem diversas áreas. Na defesa, por exemplo, as características de rápido emprego, auto-organização e tolerância a falhas vislumbram emprego em sistemas de comando, controle, inteligência, vigilância e reconhecimento. Na área médica, as RSSFs podem ser usadas para monitorar pacientes e assistir deficientes físicos. Outras aplicações incluem o monitoramento de áreas sujeitas a desastres, da qualidade da produção, das condições climáticas, dentre outras.

2.2.1 Funcionamento Básico

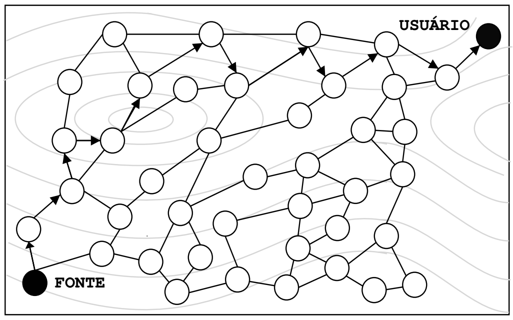
O funcionamento básico das RSSFs consiste, inicialmente, no lançamento dos sensores como um todo, ou colocados um a um na área a ser sensoriada, como ilustrado na figura 2.1(a). Em seguida, vem a fase de descoberta de rotas, onde os nós estabelecem as rotas que serão usadas para o tráfego de dados. O nó usuário então difunde as tarefas de sensoriamento na rede e aguarda as respostas. Quando um evento de interesse é detectado, os nós coletam a informação e, após algum processamento, enviam os dados ao nó usuário pelo caminho definido pelo protocolo de roteamento, como ilustrado na figura 2.1(b). A dinâmica da rede, provocada pelo movimento ou pela falha dos nós, exige uma manutenção periódica da rota a ser usada. O nó usuário possui características especiais e é responsável por coletar a informação dos nós sensores.

2.2.2 Visão Global e Principais Requisitos

As características e requisitos derivam das limitações tecnológicas e das aplicações, embora alguns destes sejam compartilhados por todas as RSSFs independente do emprego. A seguir serão apresentados os principais requisitos e características que têm impacto direto na arquitetura e nas decisões de projeto de uma RSSF e,



(a) Lançamento dos nós sensores.



(b) Estabelecimento das rotas.

Figura 2.1: Funcionamento Básico.

por conseguinte, consistem em alvo de estudos de vários pesquisadores.

A alta confiabilidade da tecnologia dos circuitos integrados e a operação em ambientes controlados restringiram a importância da **tolerância a falhas** na grande maioria dos sistemas computacionais. Entretanto, as características de emprego de uma RSSF em ambiente potencialmente hostil, com uma aplicação freqüentemente crítica e em modo contínuo, eleva em muito as chances de falhas. Os nós sensores podem falhar ou tornarem-se inoperantes por falta de energia, dano físico, ou ainda, interferência. As severas restrições de consumo de energia tornam inadequadas as técnicas tradicionais baseadas na redundância dos componentes dos nós sensores [27]. Koushanfar *et al.* [28] propõem um sistema de *backup* heterogêneo para RSSFs, no qual um tipo de recurso é substituído por outro para compensar eventuais falhas. Neste sistema, se a largura de banda para comunicação é reduzida, mas existe capacidade de processamento disponível, é possível comprimir os dados utilizando esquemas de compressão que exigem um maior poder computacional, de forma a compensar as restrições de comunicação impostas.

A suscetibilidade a falhas e a necessidade de cobertura de extensas áreas podem exigir um **elevado número de nós sensores**. O número de nós sensores empregados pode ser da ordem de centenas ou milhares e, dependendo da aplicação, chegar a milhões [29]. Dessa característica da rede, deriva um importante parâmetro de sistemas distribuídos: a densidade dos nós. De acordo com [30], a densidade μ(R) pode ser obtida pela expressão $(N\pi R^2)/$A, onde N é o número de nós sensores com raio de transmissão R, que são lançados em uma região de área A. As redes densas favorecem os algoritmos localizados e aumentam a resolução de sensoriamento.

Ao considerar o emprego das RSSFs com um elevado número de nós, o **baixo custo** de cada nó torna-se muito importante para viabilizar sua utilização. Akyildiz *et al.* [29] ressaltam que o custo de um nó sensor deveria ser inferior a um dólar americano, de forma a tornar viável o emprego de RSSFs. Os autores ressaltam que esse valor é dez vezes menor do que um rádio *Bluetooh* [31], o qual é reconhecido como um dispositivo de baixo custo.

O **tamanho físico reduzido** é um dos principais problemas enfrentados nos projetos de RSSFs. Todos os componentes do nó sensor (processador, memória, rádio, dentre outros) devem ser acomodados em um espaço de dimensões extremamente reduzidas. Os autores em [32] sugerem uma dimensão inferior a um centímetro cúbico, pesando menos do que cem gramas e um nível de dissipação de energia abaixo de cem *microwatts*.

A **topologia dinâmica** é outra característica comum em redes sem fio. No entanto, as mudanças de topologia das RSSFs não são atribuídas ao movimento dos nós, como acontece em redes *Ad Hoc* tradicionais ou nos sistemas celulares. Os nós sensores permanecem, em sua maioria, estacionários após o emprego. As mudanças na topologia ocorrem quando os nós deixam de operar por falta de energia, são destruídos, ou os rádios são desligados para economizar energia [33].

Os nós sensores comunicam-se pelo **meio de transmissão sem fio**, empregando diversas tecnologias como o rádio, o infravermelho e a óptica. A maioria do *hardware* usado em RSSFs é baseada em enlace de freqüência de rádio (RF). Por exemplo, o nó sensor μAMPS [34] utiliza *Bluetooh* [31] com transceptor operando na faixa de 2.4 GHz. Outros exemplos são o dispositivo sensor descrito por Woo e Culler [35], que usa um canal simples com transceptor RF na faixa de 916MHz, e a arquitetura das Redes de Sensores Sem Fio Integrada (*Wireless Integrated Networks Sensors - WINS*), apresentada por Pottie e Kaiser [23], que também utiliza ligação de rádio para comunicações. O emprego do infravermelho para comunicação entre os nós pode ser usada, com a vantagem da robustez a interferências elétricas e o baixo custo dos transceptores. Uma outra solução é o meio óptico, empregado no *Smart Dust* [36]. Entretanto, as tecnologias infravermelho e óptica possuem a desvantagem de precisar de uma linha de visada entre o receptor e o transmissor.

Os problemas de **segurança** em RSSFs são similares aqueles encontrados em redes *Ad Hoc* típicas. No entanto, os mecanismos de proteção desenvolvidos para redes *Ad Hoc* não são diretamente aplicáveis em RSSFs. A segurança em sentido amplo envolve os aspectos relativos a disponibilidade, a confidencialidade, a integridade e ao não repúdio de uma mensagem enviada [37], sendo necessário o desenvolvimento de soluções específicas para RSSFs como, por exemplo, o SPINS (*Security Protocols for Sensor Networks*) [38] constituído por dois protocolos: o μTESLA (*Timed, Efficient, Streaming, Loss-tolerant Authentication Protocol*) para autenticação de mensagens em difusão, e o SNEP (*Secure Network Encryption Protocol*) projetado para garantir a confidencialidade dos dados.

O problema da **localização** consiste na capacidade de um nó sensor conhecer sua posição física dentro da rede. Em algumas aplicações, os nós podem determinar suas ações com base na sua localização, ou ainda, utilizar a informação da posição para rotear os dados [39]. A solução para o problema não é simples, uma vez que uma RSSF típica consiste de um grande número de nós empregados com distribuição aleatória, não sendo possível determinar, antecipadamente, a posição de cada nó. A utilização do GPS (*Global Positioning System*) em todos os nós possui severas restrições, tais como a elevação do custo, o aumento do tamanho do nó, o maior consumo de energia e a restrição de não funcionar em ambientes fechados ou submersos. Bulusu *et al.* [30] apresentam uma técnica de localização baseada na comunicação entre pontos de referência (com posição conhecida) e nós que necessitam conhecer suas localizações.

A **sincronização** é fundamental em qualquer sistema distribuído, e particularmente, em RSSFs. A heterogeneidade dos requisitos, as restrições de consumo de energia e outras características típicas das RSSFs, e que não são encontradas em sistemas distribuídos convencionais, tornam os esquemas de sincronização existentes inadequados para este novo paradigma de comunicação. Elson [40] argumenta que os métodos atuais de sincronismo devem ser ampliados e combinados de forma a fornecerem um serviço que atenda as necessidades das aplicações das RSSFs. O autor também discute a implementação de um esquema baseado na idéia *post-facto synchronization* que, embora não possa ser aplicado em todas as situações, permite criar um "instante" de tempo sincronizado entre nós sensores.

Dentre todos os aspectos apresentados, destaca-se a característica do nó sensor ser um dispositivo microeletrônico com **fonte de energia restrita**. Na maioria das aplicações, é impossível fazer a substituição da fonte de energia. Dessa forma, as capacidades de sensoriamento, processamento e comunicação do nó sensor são limitadas pela disponibilidade de energia. A otimização do consumo de energia em RSSFs é extremamente complexa, pois não se limita a diminuir o consumo de um simples sensor, mas também a prolongar o tempo de vida da rede [41]. Para alcançar esse objetivo, é necessário que um judicioso controle de dissipação de energia seja incorporado em todos os estágios do projeto e operação de uma RSSF. O sistema deve ainda ser capaz de fazer avaliações dinâmicas entre o consumo de energia, nível de desempenho e fidelidade operacional.

2.3 Arquitetura da Rede de Sensores

Conceitualmente, uma RSSF é modelada como um conjunto de nós distribuídos em uma região onde se quer observar um determinado fenômeno, denominada de campo de sensoriamento, como ilustrado na figura 2.2. O nó sorvedouro (*sink*) é responsável pela difusão das tarefas de sensoriamento (ou interesses) e coleta dos dados sensoriados, podendo atuar também como "ponte" para redes externas. Os nós fontes são responsáveis pela monitoração do fenômeno de interesse, isto é, pela detecção das alterações físicas do ambiente (por exemplo, de temperatura, da umidade, da acústica, dentre outras) e envio dessas informações. Os nós intermediários são responsáveis pelo roteamento das informações entre os nós fonte e sorvedouro. Em princípio, todos os nós podem assumir o papel de sorvedouro, nó intermediário e nó fonte, o que garante uma grande flexibilidade na topologia da rede. Essa dinâmica de operação e a severa restrição de energia implicam em uma arquitetura de rede específica envolvendo os aspectos de *hardware*, *software* e *middleware*.

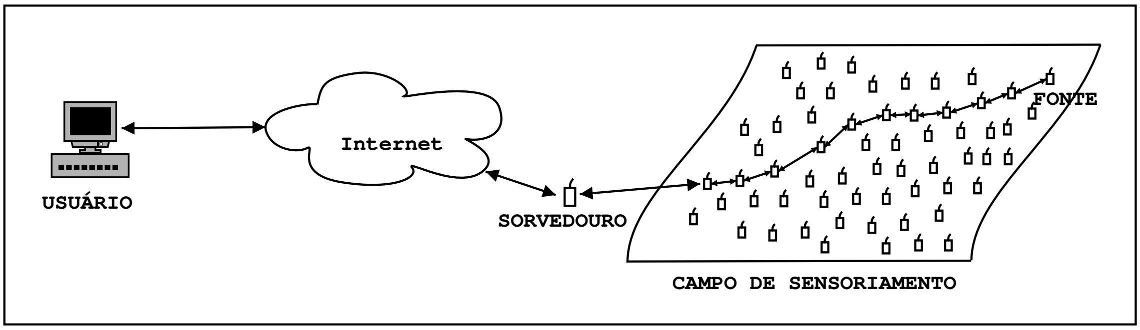


Figura 2.2: Arquitetura de rede de sensores.

Os componentes de ***hardware*** de um nó sensor típico compreendem quatro subsistemas principais: o de sensoriamento, o de processamento e de armazenamento, o de comunicação e o de energia. Dependendo da aplicação, podem existir outros subsistemas adicionais como, por exemplo, o subsistema de localização (para determinar a posição do nó com acurácia) e o subsistema de movimento (para mover o nó para um local que permita a realização de uma tarefa) [41].

O subsistema de sensoriamento é geralmente composto pelos sensores e o conversor de sinal analógico para digital (*Analogic to Digital Converter* - ADC). Os sinais analógicos produzidos pelos sensores, com base no fenômeno observado, são convertidos em sinais digitais pelo ADC, para serem tratados pelo processador. Existe uma variedade de sensores capazes de gerar uma resposta mensurável para parâmetros ambientais, tais como de temperatura, de luminosidade, de acústica, etc. Com relação ao consumo de energia, no sensor passivo (por exemplo, o termômetro) pode ser considerado desprezível comparado com outros componentes do nó, enquanto no sensor ativo (por exemplo, o sonar) não pode ser negligenciado.

O subsistema de processamento, que geralmente está associado à unidade de armazenamento local, é responsável pela execução dos protocolos de comunicação e dos algoritmos de processamento, pelo controle dos sensores, e ainda, pela gerência dos procedimentos para que os nós sensores trabalhem de forma colaborativa. A demanda de desempenho exigida pela aplicação é diretamente proporcional ao consumo de energia do subsistema.

Um rádio de curto alcance compõe o subsistema de comunicação. A dissipação de energia do rádio depende de diversos fatores, tais como: o tipo de modulação, a taxa de transmissão, a potência utilizada e o ciclo de serviço. A comunicação via rádio responde pela maior parcela do consumo de energia quando comparada ao sensoriamento ou ao processamento. Dessa forma, é desejável que a comunicação ocorra a curtas distâncias, utilizando nós intermediários em rotas com múltiplos saltos.

A fonte de energia engloba uma bateria e um conversor DC-DC. A bateria é um dispositivo complexo cuja operação depende de diversas características como, por exemplo, as dimensões, o material usado e a taxa de descarga. O conversor DC-DC fornece uma tensão constante para o nó sensor e o seu fator de eficiência determina o tempo de vida da bateria. Em [27] os autores discutem que, conceitualmente, existem no mínimo duas opções para fonte de energia em RSSFs. A opção dominante considera o uso de células de bateria de alta densidade. A outra alternativa são as células solares, empregadas em áreas com luminosidade capaz de atender a demanda do nó sensor.

Com relação ao ***software***, a principal preocupação continua sendo encontrar soluções que garantam um baixo e bem distribuído consumo de energia na rede. Nesse aspecto, o sistema operacional assume importância fundamental para facilitar a comunicação entre os nós através de múltiplos saltos, a necessidade de privacidade e segurança, e a execução eficiente de algoritmos (localizados) específicos da aplicação. Um exemplo de sistema operacional proposto especificamente para RSSF é o *TinyOS* [27]. O *TinyOS* é um sistema operacional modular baseado em eventos, que ocupa pouca memória e possui um baixo consumo de energia, sendo projetado para uma intensiva operação concorrente e robusta [42].

O ***middleware*** atua como intermediário entre as aplicações e o ambiente operacional da RSSF, o que facilita o desenvolvimento de novas aplicações como, por exemplo, a agregação de dados [27]. Shen *et al.* [43] introduzem um *middleware* denominado SINA (*Sensor Information Networking Architecture and Applications*) que permite que as aplicações façam consultas, designem tarefas, coletem respostas e resultados, e monitorem as modificações na rede, utilizando a linguagem SQTL (*Sensor Query and Tasking Language*) como API (*Application Programming Interface*).

2.4 Pilha de Protocolos

A pilha de protocolos usada pelo sorverdouro e os nós sensores caracteriza-se por explorar o esforço cooperativo e a disponibilidade de energia dos elementos da rede para o estabelecimento das rotas, a agregação de dados e a comunicação empregando o meio sem fio. Esta seção aborda alguns aspectos referentes a pilha de protocolos da camada física, de enlace, de rede, de transporte e de aplicação usada pelas RSSFs [29], destacando os principais trabalhos relacionados.

2.4.1 Camada Física

Segundo o modelo de propagação de rádio de primeira ordem apresentado em [44], a comunicação entre os nós (transmissão e recepção) é responsável pelo maior dispêndio de energia em uma RSSF. Portanto, a eficiência em termos de consumo de energia assume uma importância significativa no projeto da camada física, influenciando os estudos sobre a seleção da freqüência, da portadora, detecção do sinal e modulação [29].

A faixa de freqüência ISM (*Industrial, Scientific, and Medical*) tem sido amplamente sugerida para RSSFs [29]. Por exemplo, no projeto WINS (*Wireless Integrated Network Sensors*) [23] é utilizada a comunicação por freqüência de rádio (RF) nas faixas 902-928 MHz e próximo de 2.4GHz. Estas faixas de freqüência exploram favoravelmente o compromisso entre o consumo de energia associado com a transmissão e as dimensões das antenas [45].

Uma abordagem empregada para reduzir a energia gasta com a comunicação é diminuir o tempo de transmissão do rádio, o que pode ser obtido enviando múltiplos *bits* por símbolo, através de uma modulação M-ária. No entanto, o trabalho de Shih *et al.* [34] mostra que a complexidade dos circuitos pode elevar o consumo total de energia do rádio quando comparada a modulação binária.

Chien *et al.* [46] publicaram um estudo propondo uma arquitetura de espalhamento do espectro por seqüência direta para redes de sensores com consumo mínimo de energia. Essa arquitetura foi implementada em FPGA (*Field Programmable Gate Array*) e verificado que a dissipação de energia, tanto na transmissão como na recepção, é da ordem de 33mW. Os autores estimam um desempenho menor que 1mW, empregando tecnologia ASIC (*Application-specific Integrated Circuit*).

O rádio impulso (IR) ou *Ultra Wideband* (UWB) tem despertado interesse para aplicações de RSSF em ambientes fechados [47]. O IR emprega transmissão em banda base e, geralmente, modulação PPM (*Pulse Position Modulation*), sendo que o baixo consumo de energia, a simplicidade do circuito e a capacidade de detectar e ignorar sinais refletidos em decorrência de múltiplos caminhos, tornam o seu emprego bastante atrativo [29].

As redes *Bluetooth* [31] e 802.15.4/ZigBee [48] representam as tecnologias atuais mais próximas das RSSFs. *Bluetooth* é uma tecnologia para comunicações de curto alcance, com baixo consumo de energia, baixo custo e direcionada a substituir os cabos de conexão dos dispositivos eletrônicos por enlaces de rádio. A estrutura básica do *Bluetooth* é a picorede, formada por um nó mestre e até sete dispositivos (escravos), formando uma topologia em estrela. A comunicação ocorre através de TDMA e emprega a técnica de salto de freqüências. A tecnologia *Bluetooth* foi projetada para atender uma variedade de dispositivos, o que pode provocar desperdício de energia em alguns casos. Esse problema é resolvido com as redes 802.15.4/ZigBee, que permitem conectar de forma mais eficiente dispositivos que requerem baixo consumo de energia. As redes 802.15.4/ZigBee podem ter até 250 dispositivos, operam na faixa de freqüência ISM de 2.4 GHz, empregando CSMA-CA e a técnica de espalhamento do espectro por seqüência direta.

2.4.2 Camada de Enlace

O objetivo da camada de enlace é assegurar conexões confiáveis em uma rede de comunicação. Isto é realizado através das tarefas de multiplexação dos fluxos de dados, detecção dos quadros, acesso ao meio e o controle de erro. Embora muitas dessas funções ainda não tenham sido completamente definidas, o protocolo de acesso ao meio (*Medium Access Control* - MAC) é o mais crítico para diversas aplicações de RSSFs.

Como em toda rede que compartilha o meio de transmissão, o protocolo MAC é fundamental para o sucesso da operação da rede. A função mais importante do protocolo é evitar colisões. Existem diversos protocolos MAC disponíveis para redes sem fio. No entanto, esses protocolos possuem restrições quando aplicados às RSSFs [29]. Por exemplo, o acesso múltiplo por divisão do tempo (*Time Division Multiple Access* - TDMA) exige uma elevada taxa de transmissão e coordenação entre os nós, enquanto o acesso múltiplo por divisão de código (*Code Division Multiple Access* - CDMA) e os protocolos baseados em contenção da especificação IEEE 802.11 requerem que os nós escutem o canal para receber um possível tráfego, o que provoca um elevado dispêndio de energia. Tais aspectos implicam em modificações ou desenvolvimento de novos protocolos de acesso ao meio, específicos para RSSFs.

No estudo de Ye *et al.* [49] são apresentados os requisitos do protocolo MAC de uma RSSF e é proposto o S-MAC (*Sensor MAC*), cujo o principal objetivo é reduzir o consumo de energia evitando o desperdício com a colisão, o recebimento de pacotes de destinatário diferente, as mensagens de controle e o rádio aguardando para receber dados. O S-MAC é baseado em um esquema de contenção para evitar as colisões, no qual os nós escutam o meio periodicamente, desligando o rádio nos intervalos.

O SMACS (*Self-Organizing Medium Access Control for Sensor Networks)* [22] é um protocolo que permite a descoberta de nós vizinhos e o agendamento da comunicação entre eles, sem a necessidade de um nó de coordenação. O protocolo combina em uma única fase a descoberta dos nós vizinhos e a designação do canal para o enlace. Cada enlace estabelecido opera com uma freqüência diferente para evitar colisões. Dessa forma, o nó sabe quando deve ligar o transceptor para realizar uma comunicação, mantendo-o desligado quando não existe nenhuma transmissão ou recepção agendada.

Estudos adicionais para controle de acesso ao meio em RSSFs incluem modificações no protocolo tradicional de múltiplo acesso com detecção da portadora (*Carrier Sense Multiple Access* - CSMA), proposto por Woo *et al.* [35], e um esquema híbrido baseado em TDMA/FDMA, introduzido por Shih *et al.* [34].

Outra importante função da camada de enlace é o controle de erro na transmissão de dados. Os dois modos mais importantes de controle de erro em redes de comunicações são o ARQ (*Automatic Repeat Request*) e o FEC (*Forward Error Correction*). O ARQ é inapropriado porque as retransmissões exigem um custo extra de energia [29], enquanto o FEC eleva exponencialmente o consumo médio de energia (por *bit*) com o aumento do comprimento do código [34]. Esses aspectos sugerem o desenvolvimento de protocolos de controle de erro simples com baixa complexidade na codificação e decodificação ou modificação nos existentes para obtenção de uma maior eficiência em RSSFs.

2.4.3 Camada de Rede

Os protocolos da camada de rede são responsáveis pelo roteamento dos dados. No entanto, ao contrário dos protocolos tradicionais que buscam diminuir o retardo fim-a-fim ou aumentar a vazão, em RSSFs eles visam estabelecer rotas que permitam estender o tempo de vida da rede por meio da racionalização do consumo de energia, com sacrifício de outras métricas de desempenho [50]. Rotas eficientes em termos de consumo de energia são baseadas na disponibilidade ou demanda necessária para transmissão. Existem diferentes abordagens para obter um protocolo de roteamento eficiente em energia. Por exemplo, a rota escolhida para transmissão pode ser baseada na máxima energia disponível, na mínima energia consumida, no menor número de transmissões (saltos) ou, ainda, em um modo híbrido das abordagens anteriores. Além da conservação de energia, outras características importantes destes protocolos são o roteamento centrado em dados e a agregação.

No roteamento centrado em dados, as rotas são definidas com base nos interesses enviados pelo sorvedouro e nos dados coletados pelas fontes. Existem dois esquemas utilizados para disseminar interesses: no primeiro, o sorvedouro faz a difusão dos interesses [7] e no segundo, os nós sensores difundem a informação sobre a disponibilidade dos dados e esperam uma requisição com interesses [51]. Este tipo de roteamento adota um sistema de endereçamento baseado em atributos [43], no qual os usuários estão mais interessados em obter informação a respeito de um determinado fenômeno do que consultar um nó individualmente.

O trabalho proposto por Park *et al.* [52] ressalta que, ao invés de enviar dados brutos (provavelmente comprimidos) para o sorvedouro, os nós podem realizar um processamento na rede para tomada de decisão. Tal processamento é definido como agregação de dados. Na agregação, uma RSSF é vista como uma árvore ($N \mapsto 1$), onde o sorvedouro solicita aos nós sensores que relatem as condições de um determinado fenômeno. Os dados que chegam de múltiplos nós sensores são agregados a cada nó com base nos atributos observados. Heinzelman *et al.* [51] explicam a agregação de dados como uma técnica empregada para solucionar o problema de implosão (quando mensagens duplicadas são enviadas para o mesmo nó) e da sobreposição (quando nós distintos percebem o mesmo fenômeno e geram mensagens duplicadas) no roteamento centrado em dados. O capítulo 3 irá apresentar, com detalhes, a agregação de dados empregando técnicas paramétricas, que foram implementadas no mecanismo proposto nesta dissertação.

Os protocolos de roteamento de redes *Ad Hoc* estão organizados em duas categorias: *table-driven* e *demand-driven* [53]. A primeira, requer que cada nó mantenha uma ou mais tabelas com a informação das rotas - é o caso do DSDV (*Destination-Sequenced Distance-Vector Routing*) e o WRP (*Wireless Routing Protocol*), enquanto a segunda, utiliza um processo de descoberta de rotas quando é necessário conhecer um caminho para um nó destino - é o caso do AODV (*Ad Hoc On-Demand Distance Vector Routing*) e o DSR (*Dynamic Source Routing*). Estas abordagens de roteamento de redes *Ad Hoc* tradicionais não são apropriados para as RSSFs, devido às características apresentadas na seção 2.2.2, havendo portanto, a necessidade de desenvolvimento de novos protocolos. A seguir serão abordados diferentes propostas de algoritmos de roteamento em RSSF, dentre os mais referenciados na literatura e que, atualmente, servem de base para a maioria das pesquisas.

A **inundação** (*flooding*) é uma técnica antiga que pode ser usada para roteamento de dados em RSSF. Na inundação, cada nó que recebe uma mensagem se encarrega de difundí-la, exceto quando o número máximo de saltos é alcançado ou o nó é o destinatário. A inundação é uma técnica reativa, não requer manutenção com a topologia e algoritmos complexos para descoberta de rotas. No entanto, existem várias deficiências, tais como: a implosão, a sobreposição e a desconsideração dos recursos de energia disponíveis. Existe ainda, uma derivação do protocolo de inundação denominada *gossiping* [29]. Em tal protocolo, o nó sensor seleciona aleatoriamente um de seus vizinhos para enviar os dados. O nó vizinho que recebeu os dados repete a operação, escolhendo ao acaso um outro nó sensor. Embora essa abordagem evite a implosão (devido ao fato de que só existe uma cópia da mensagem em cada nó), pode tomar um tempo muito longo para propagar a mensagem.

Rodoplu e Meng [54] criaram um protocolo denominado MECN (*Minimum Energy Communication Network*) que, a partir de uma rede de comunicações, calcula uma subrede tal que, para cada par (u,v) de nós conectados na rede original, existe um caminho entre u e v na subrede, que garante o uso mínimo de energia para transmissão de mensagens entre os dois nós. Com base no protocolo MECN, Li e Halpern [55] construíram o protocolo **SMECN** (*Small Minimum Energy Communication Network Revisited*) que utiliza os mesmos princípios, mas com uma subrede menor. Os resultados do SMECN mostraram uma significante economia de energia e uma maior simplicidade computacional.

Uma família de protocolos adaptativos denominados **SPIN** (*Sensor Protocols for Information via Negotiation*) foi proposto por Heinzelman *et al.* [51] para resolver as deficiências do protocolo de inundação. O SPIN dissemina a informação eficientemente via negociação, de forma a garantir uma maior conservação de energia.

O protocolo possui três tipos de mensagens básicas: de aviso (ADV), de requisição (REQ) e de dados (DATA). Inicialmente, o sensor difunde uma mensagem de aviso contendo um metadado (isto é, um descritor dos dados). Se o vizinho está interessado nos dados, ele envia uma mensagem de requisição e somente então os dados serão enviados. O processo é repetido até que todos os nós sensores da rede, que tenham interesse nos dados, recebam uma cópia.

O algoritmo **SAR** (*Sequential Assignment Routing*) faz parte de um conjunto de algoritmos definidos por Sohrabi *et al.* [22] para a organização, gerenciamento e mobilidade em uma RSSF. Os outros algoritmos são o SMACS, visto na seção 2.4.2 e o EAR (*Eavesdrop-And-Register*), que foi projetado para permitir uma interconexão contínua de nós móveis com nós estacionários, baseando-se em mensagens convite e registro. O SAR gera múltiplas árvores onde o nó raiz está a um salto do sorvedouro. Cada árvore é construída evitando-se nós com baixa QoS (*Quality of Service*) e disponibilidade de energia. Ao final do procedimento, a maioria dos nós irá pertencer a múltiplas árvores, o que permitirá que cada nó selecione o seu caminho com base no parâmetros de QoS, energia e prioridade do pacote. Integram ainda esse conjunto de algoritmos o SWR (*Single Winner Election*) e o MWE (*Multi Winner Election*), que são responsáveis pela sinalização e transferência de dados.

O protocolo **LEACH** (*Low-Energy Adaptive Clustering Hierarchy*) é baseado no agrupamento de nós com objetivo de minimizar a dissipação de energia [44]. O LEACH seleciona aleatoriamente um nó do agrupamento (*cluster-head*) para ser responsável pela agregação dos dados dos sensores e envio para o sorvedouro. O protocolo original emprega uma rotação não determinística do *cluster-head* para distribuir a carga de energia na rede. A escolha aleatória do *cluster-head* foi modificada em [56], por um algoritmo determinístico, cujos resultados dessa alteração mostraram um aumento do tempo de vida da rede da ordem de 30%. Outra evolução do LEACH foi a introdução do protocolo PEGASIS (*Power-Efficient Gathering in Sensor Information Systems*) [57]. A idéia chave do protocolo é formar uma cadeia de nós, onde cada nó só se comunica com o vizinho mais próximo e, a cada turno, um nó é designado aleatoriamente para transmitir os dados agregados para a estação base. O resultado da simulação do PEGASIS mostra um desempenho superior ao

LEACH, com a desvantagem do aumento da latência da rede.

O protocolo de **Difusão Direcionada** (*Directed Diffusion*) foi proposto por Intanagonwiwat *et al.* [7]. Trata-se de um protocolo centrado em dados, no qual o sorvedouro envia interesses para todos os sensores da rede e quando esses dados se tornam disponíveis nos nós fontes, eles são roteados por múltiplos caminhos para o nó que originou o interesse. Cabe ao sorvedouro reforçar um nó vizinho em particular, segundo uma métrica estabelecida. Os nós intermediários usam o mesmo critério para reforçar seus respectivos vizinhos, criando uma rota até a fonte. O protocolo de Difusão Direcionada será visto com maiores detalhes na seção 3.1.2.

2.4.4 Camada de Transporte

A camada de transporte assume uma importância fundamental quando a RSSF é planejada para ter acesso a redes externas como, por exemplo, a Internet. Tanenbaum [13] discute em seu livro os problemas de desempenho decorrentes da implementação do TCP (*Transmission Control Protocol*) em redes de comunicação sem fio e apresenta o TCP indireto como uma possível solução. Em uma RSSF, essa solução resultaria na comunicação entre o usuário e o sorvedouro usando TCP ou UDP (*User Datagram Protocol*) via Internet, e entre o sorvedouro e os nós sensores seria usado o UDP puro. Entretanto, alguns autores [29] argumentam que as características específicas das RSSFs justificam a necessidade de desenvolvimento de novos protocolos para a camada de transporte.

O protocolo ESRT (*Event-to-Sink Reliable Transport in Wireless Sensor Networks*) foi introduzido por Yogesh *et al.* [58] como uma nova solução para camada de transporte em RSSFs. A comunicação fim-a-fim é determinada pelo estado corrente da rede, que é baseado no nível de confiabilidade e condição de congestionamento. O ESRT define cinco estados possíveis: sem congestionamento e baixa confiabilidade, sem congestionamento e alta confiabilidade, congestionado e com alta confiabilidade, congestionado e baixa confiabilidade, e operação ótima. O objetivo primário do ESRT é alcançar e manter a rede na região ótima, onde a confiabilidade necessária é atingida com o mínimo consumo de energia e sem congestionamento da rede.

Outro protocolo desenvolvido para explorar a confiabilidade na camada de transporte de redes de sensores foi apresentado por Stann e Heidemann [59]. O protocolo RMST (*Reliable Multi-Segment Transport*) foi projetado para operar em conjunto com a Difusão Direcionada. A finalidade do protocolo é garantir a entrega de todos os dados solicitados pelo sorvedouro, que sejam oriundos de uma mesma fonte. O RMST realiza a fragmentação e a remontagem das mensagens, adicionando atributos específicos aos dados que fluem de uma fonte para o sorvedouro, com um custo mínimo de controle. Os autores concluem que o RMST constitui uma base para expansão das aplicações da Difusão Direcionada que requerem confiabilidade.

2.4.5 Camada de Aplicação

O amplo potencial das aplicações que empregam a tecnologia de RSSFs tem motivado consideravelmente a pesquisa na área. As aplicações estendem-se por diversos segmentos do conhecimento humano. Na medicina existem estudos de sensores inteligentes para correção da visão [60] e monitoramento de pacientes [55]. Com relação ao meio ambiente, uma RSSF pode ser empregada para monitorar um *habitat* específico, com a finalidade de identificar, gravar e analisar eventos de interesse [61], ou ainda, ser lançada em uma área de desastre (por exemplo, um incêndio) para auxiliar em um possível resgate [62]. Na área de defesa, as principais aplicações visam detectar, identificar e rastrear alvos inimigos utilizando menos proteção blindada [63]. Aplicações na área de segurança tem por objetivo proteger instalações contra ameaças [23]. Outras propostas para aplicações em RSSFs incluem o monitoramento de linha de produção [64], o controle de inventário de fábrica e o controle de tráfego em áreas urbanas [62].

Apesar das diversas propostas de aplicações, Akyildiz *et al.* [29] afirmam em sua pesquisa sobre RSSFs que os protocolos para camada de aplicação ainda são pouco explorados e sugerem três possíveis protocolos de aplicações de apoio: o de gerenciamento de sensores; o de aviso de dados e designação de tarefas; e o de consultas e disseminação de dados. O protocolo de gerenciamento de sensores tem por finalidade tornar o *hardware* e o *software* das camadas mais baixas transparentes para as aplicações que gerenciam a rede. O protocolo deve acessar os nós (que não possuem identificadores globais) para realizarem tarefas administrativas como, por exemplo, a manutenção nas regras de agregação, a sincronização, a alteração da posição, a autenticação e ainda, ligar e desligar o sensor. O protocolo de aviso de dados e designação de tarefas deve gerenciar a sinalização do sensor que possui disponibilidade de dados e a disseminação de interesses dos usuários. Uma interface eficiente para estas operações auxilia as camadas mais baixas como, por exemplo, no roteamento. O protocolo de consulta e disseminação de dados estabelece uma interface para às aplicações dos usuários consultarem, responderem e receberem respostas da RSSF.

2.5 Conclusão Parcial

O presente capítulo apresentou os principais aspectos envolvidos em RSSFs, incluindo o funcionamento básico, as características, os requisitos comuns e a arquitetura da rede no que se refere ao *hardware*, *software* e *middleware*. Também foram apresentadas as questões relacionadas a pilha de protocolos usada pelos nós sensores para comunicação.

Das questões apresentadas é possível concluir que a energia constitui o fator mais importante de uma RSSF. O tempo de vida da rede depende da energia economizada em cada solução adotada para os protocolos da camada física, de enlace, de rede, de transporte e de aplicação, sendo os maiores ganhos obtidos com a redução da comunicação entre os nós. O capítulo a seguir dedica-se ao estudo de dois importantes aspectos para a obtenção de uma maior eficiência em termos de energia: o protocolo de roteamento e a agregação de dados.

Capítulo 3

Roteamento e Agregação de Dados

Em RSSFs, o sensoriamento de um fenômeno pode gerar um elevado nível de transmissões de dados, o que compromete o tempo de vida da rede. Essa característica implica a necessidade de um processamento distribuído na rede para reduzir a comunicação entre os nós. Esse tipo de processamento é realizado pela agregação de dados enquanto estes são roteados, fazendo uso de um modelo de roteamento centrado em dados, em oposição à abordagem tradicional centrada em endereços. As técnicas de agregação empregadas podem realizar desde uma simples supressão de mensagens até a aplicação dos métodos de fusão de dados preconizados pelo modelo JDL DFS como, por exemplo, as técnicas paramétricas. A Difusão Direcionada é um protocolo de roteamento centrado em dados que favorece o emprego destas técnicas.

Neste capítulo são apresentadas as questões relativas aos modelos de roteamento, explorando em detalhes o protocolo de roteamento centrado em dados da Difusão Direcionada e as técnicas paramétricas de fusão de dados, objetos de estudo do trabalho descrito nesta dissertação. A seção 3.1 apresenta os modelos de roteamento e o procedimento para formação das rotas empregando o protocolo da Difusão Direcionada. A seção 3.2 apresenta o conceito de agregação de dados, sua finalidade e seus impactos em uma RSSF. Nessa seção são exploradas, ainda, as técnicas da Inferência Bayesiana e de Dempster-Shafer para agregação de dados. Estes assuntos foram abordados em detalhes por estarem diretamente relacionados aos objetivos desta dissertação. Finalmente, a seção 3.3 traz as conclusões do capítulo.

3.1 Roteamento

A comunicação entre os nós é a maior responsável pelo dispêndio de energia em RSSFs, como foi visto na seção 2.4.1. Pottie *et al* [23] publicaram um estudo concluindo que o custo da comunicação é muito superior ao custo do processamento em cada nó. Com um modelo de comunicações empregando uma portadora de 1 GHz, modulação BPSK (*Binary Phase Shift Keying*) com probabilidade de erro de 10^{-6}, desvanecimento de *Rayleigh* e um receptor ideal, os autores mostram que o custo de se transmitir 1Kb na distância de 100 metros é de aproximadamente de 3 *joules*. Em contraste, um processador genérico com 100MIPS/W pode executar 300 milhões de instruções com a mesma quantidade de energia. Essa característica motiva o uso da agregação de forma distribuída, enquanto os dados são roteados entre os nós fontes e o sorvedouro. Para que a *agregação-na-rede* seja realizada, um esquema apropriado para o endereçamento e encaminhamento dos dados das fontes até o sorvedouro é fundamental.

3.1.1 Modelos de Roteamento

Segundo Krishnamachari *et al.* [3], a forma como as rotas entre as fontes e o sorvedouro são estabelecidas permite classificar dois tipos de protocolos de roteamento: o protocolo centrado em endereços, em que cada fonte envia dados de forma independente para o destino; e o protocolo centrado em dados, em que as rotas são definidas com base no conteúdo das mensagens. A pesquisa dos autores mostrou que os protocolos centrados em endereços não possuem a mesma eficiência dos protocolos centrados em dados ao explorar o nível de redundância não determinístico dos dados em RSSFs. Dessa forma, os autores concluem que o protocolo centrado em dados é a melhor opção para a implementação efetiva de uma forma de agregação de dados de múltiplas fontes.

No roteamento centrado em dados, a designação de um identificador global único para cada nó, estilo endereço IP (*Internet Protocol*), é substituída por um endereçamento baseado em atributos. Um exemplo que explora a comunicação centrada em dados é o *Data Funneling* [65], no qual apenas um fluxo de dados de um grupo de fontes é enviado até o destino. O algoritmo combina agregação com compressão de dados em um campo de sensores heterogêneos, no qual nós controladores com maior capacidade de processamento são responsáveis pela divisão da área em regiões distintas e pela disseminação dos interesses em cada uma dessas regiões. Ao atingir a borda da região alvo, o interesse é modificado para que os nós internos se organizem de forma sincronizada para reportar-se ao nó da borda, com o intuito de permitir o processamento dos dados, antes do seu envio em direção ao sorvedouro.

Outros exemplos incluem protocolos que buscam estabelecer árvores de agregação na rede. A formação de uma árvore de agregação de dados ótima em termos de número de transmissões é, em geral, um problema NP-completo [66]. Krishnamachari *et al.* [3] apresentam três heurísticas para geração de árvores de agregação "quase ótimas", que são baseadas no posicionamento dos nós (*Center at Nearest Source* - CNS e *Shortest Paths Tree* - SPT) e no estabelecimento de rotas consecutivas (*Greedy Incremental Tree* - GIT). A análise dos autores mostrou que a maior economia de energia ocorre para um elevado número de fontes, posicionadas mais próximas umas das outras e afastadas do sorvedouro. Intanagonwiwat *et al.* [4] complementam a pesquisa ao mostrarem a importância da energia gasta no estabelecimento e manutenção das árvores de agregação, em redes altamente densas, para obtenção de uma maior economia de energia.

Heinzelman *et al.* analisam em seu trabalho [44] o roteamento baseado em agrupamentos hierárquicos (*clusters*) que favorece a agregação nos líderes de cada *cluster* (*cluster heads*). No entanto, tal abordagem envolve uma certa complexidade e comunicação adicional para a eleição dos *cluster heads*, consumindo nesse processo uma quantidade não desprezível de energia.

A agregação de dados baseada em consultas utilizando uma linguagem estilo SQL (*Structured Query Language*) é proposta em [67]. Os autores discutem a implementação de cinco funções básicas de banco de dados (*COUNT, MIN, MAX, SUM* e *AVERAGE*) na plataforma TinyOS. Por exemplo, a declaração *SELECT AVERAGE (temperature) FROM sensors*, calcula a temperatura média das leituras realizadas pelo sensor e que estão armazenadas em uma tabela local. O esquema consiste em propagar a consulta na rede e, após um intervalo de tempo, cada nó calcula a função de agregação com base nas informações locais e nas recebidas pelos nós filhos, se houver. Essa abordagem favorece a agregação distribuída na rede, no entanto exige mais memória disponível no nó sensor e algum tipo de sincronização. A limitação das cinco funções básicas originais foram ampliadas com a apresentação do esquema TAG (*Tiny AGgregation*) [68].

O protocolo de roteamento amplamente difundido em RSSFs e que favorece o emprego de esquemas de agregação é a Difusão Direcionada. Este protocolo de roteamento é explorado, em detalhes, nesta dissertação por ter sido utilizado na implementação do mecanismo de agregação apresentado no capítulo 4. As principais características da Difusão Direcionada incluem o elevado grau de robustez, a relativa simplicidade e a independência da topologia da rede. Trata-se de um protocolo orientado a eventos, no qual os nós transmitem dados ao detectarem a ocorrência de um fenômeno de interesse. Essa característica torna o protocolo de Difusão Direcionada apropriado para o emprego de técnicas paramétricas de agregação de dados. Os principais aspectos que compõem a Difusão Direcionada [8] são abordados na seção seguinte.

3.1.2 Difusão Direcionada

Em uma rede de sensores baseada na Difusão Direcionada, os nós operam próximos ao fenômeno de interesse e de forma cooperativa para armazenar e processar as mensagens. Esta cooperação entre os nós ocorre através das interações localizadas (ou seja, mensagens trocadas entre nós vizinhos), que são responsáveis por determinar a propagação e a agregação dos dados.

O protocolo da Difusão Direcionada é o resultado de um trabalho cooperativo que envolve pesquisadores da *University of Southern California/Information Sciences Institute, University of California, BAE Systems, Cornell University* e *Pennsylvamia State University*. No estágio atual, encontra-se implementado em diversas plataformas, como por exemplo, WINSng 2.0 nodes [23], USC/ISI PC/104 nodes [69], Motes [70] e um módulo no simulador *ns-2* [9], o qual foi utilizado neste trabalho de dissertação. Uma descrição dos principais elementos do protocolo será apresentada a seguir.

Esquema de Nomes

Na Difusão Direcionada, as tarefas de sensoriamento são nomeadas por uma lista de atributos associados a valores. Estes pares atributo-valor especificam um interesse por dados cujo conteúdo tenha os mesmos atributos. Por esta razão, a descrição da tarefa é chamada de interesse. Por exemplo, uma tarefa simples para detectar um alvo é ilustrada em 3.1(a). A tarefa define o interesse em detectar um carro blindado, o intervalo de envio de mensagens, o período e a área de interesse. Os dados enviados em resposta aos interesses também seguem o mesmo esquema de nomes. Como ilustrado na figura 3.1(b), um sensor que detecta um carro blindado poderia gerar mensagens com os atributos da classe do blindado, sua localização, a medida da amplitude do sinal detectado (intensidade), o grau de confiança na detecção e o horário em que a mensagem foi gerada.

```
tipo = carro blindado
intervalo = 20 ms
duração = 10 segundos
região = [-100,100,400,400]
```

(a) Tarefa.

```
tipo = carro blindado
classe = M108
localização = [200,200]
intensidade = 0,8
confiabilidade = 0,85
tempo = 01:22:50
```

(b) Dado.

Figura 3.1: Esquema de nomes.

No exemplo apresentado, foram escolhidos valores simples para os atributos dos interesses e dos dados, os quais atendem aos objetivos desta dissertação. No entanto, uma visão mais abrangente e as questões que envolvem a escolha de um esquema de nomes podem ser encontradas no trabalho de Adjie-Winoto *et al.* [71], denominado *Intentional Naming System* (INS).

Interesses e Gradientes

Para cada tarefa de sensoriamento, o sorvedouro difunde periodicamente uma mensagem contendo um interesse. A mensagem inicial de interesse tem como objetivo determinar se existe algum sensor que tenha detectado o evento que se quer monitorar. O envio periódico do interesse é necessário devido à falta de confiabilidade da transmissão através da rede. A taxa de atualização dos interesses é um parâmetro do protocolo e explora o compromisso entre o custo adicional de enviar mais mensagens e o aumento da robustez contra perdas eventuais de pacotes. Cada nó sensor armazena uma lista de interesses distintos, como ilustrado na figura 3.2. A mensagem de interesse armazenada no nó sensor contém vários campos, tais como o horário em que o nó recebeu o interesse (ou sua atualização) e os gradientes (um para cada nó vizinho interessado). Cada gradiente possui uma taxa de envio de dados (derivada do atributo intervalo) e a duração (tempo de vida) do interesse. O interesse não possui nenhuma informação a respeito do sorvedouro, apenas sobre o nó vizinho que enviou a mensagem.

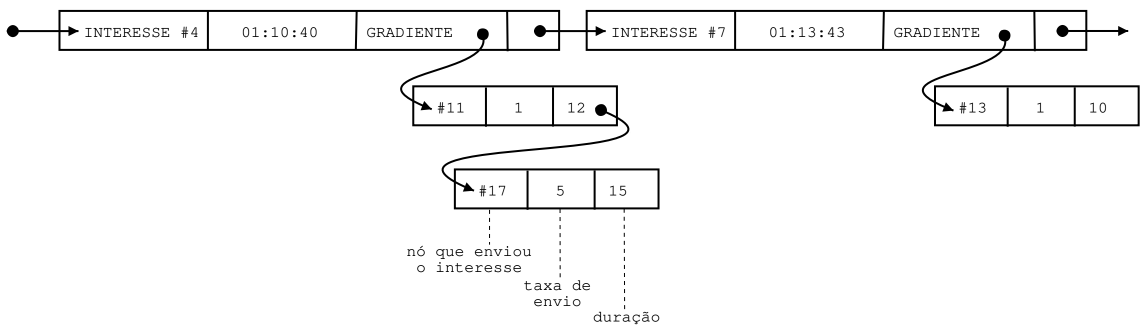
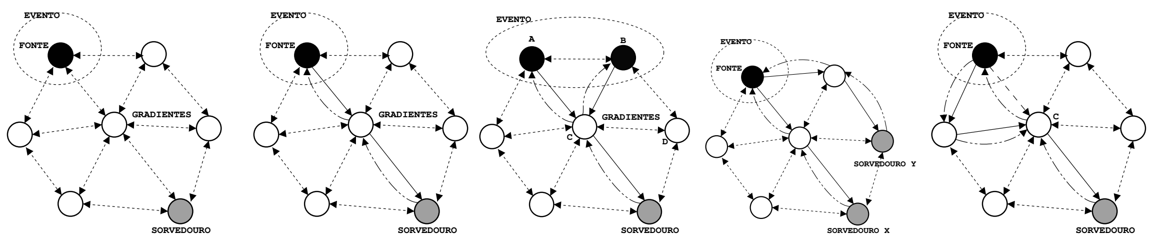


Figura 3.2: Lista de interesses.

Quando um nó recebe um interesse, ele verifica se o mesmo faz parte da sua lista. Caso o interesse não exista, ele é adicionado à lista com o respectivo gradiente direcionado para o nó que enviou a mensagem. Portanto, é necessário que o nó seja capaz de distinguir cada um de seus vizinhos. Qualquer identificador local (único para cada vizinho) pode ser utilizado, como por exemplo, o endereço MAC 802.11 [15]. No caso de existir o interesse na lista, porém sem um gradiente para o nó que enviou a mensagem, este é adicionado e o horário de recebimento do interesse é atualizado. Finalmente, quando a lista já possui o interesse e um gradiente para o nó que enviou a mensagem, somente são atualizados o horário de chegada e a duração do interesse. Após atualizar sua lista, o nó poderá enviar o interesse para alguns de seus vizinhos (podem ser excluídos aqueles que receberam interesses semelhantes recentemente). A repetição deste procedimento permite que o interesse seja disseminado pela rede.

Como conseqüência das interações localizadas, os nós que recebem um interesse consideram que a mensagem teve origem no nó que a enviou, embora a procedência da mensagem possa ser de um sorvedouro distante. Essa característica faz com que cada par de nós vizinhos estabeleça um gradiente entre si, como ilustrado na figura 3.3(a). A existência desses gradientes pode causar o recebimento de mensagens duplicadas. No entanto, esta técnica permite a seleção de caminhos empiricamente melhores e o restabelecimento rápido do fluxo de dados no caso de falhas nas rotas.



(a) Gradientes. (b) Reforço. (c) Múltiplas fontes. (d) Múltiplos sorvedouros. (e) Reparo.

Figura 3.3: Aspectos da Difusão Direcionada.

Fluxo de Dados

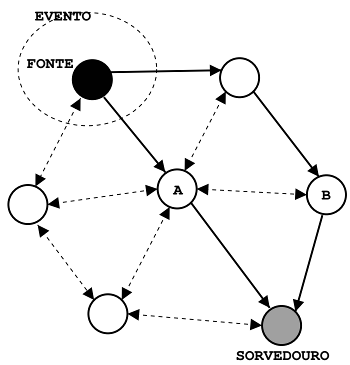
Quando um nó fonte observa um evento, ele verifica na lista de interesses se existe alguma demanda para aquele evento. Se houver, o nó seleciona a maior taxa de envio de dados dentre todos os gradientes daquele interesse, e ajusta o sensor para gerar amostras do evento nesta taxa. Em seguida, passa a enviar uma mensagem de dados (como por exemplo, a que foi ilustrada na figura 3.1(b)) para todos os vizinhos com os quais tenha um gradiente. O nó intermediário que recebe uma mensagem de dados de seu vizinho, busca um interesse em sua lista para aquela mensagem. Caso não exista, esta é simplesmente descartada. Entretanto, se houver um interesse, o nó verifica uma outra lista que mantém um histórico das mensagens de dados que foram recentemente enviadas, para evitar a formação de laços. Se a mensagem recebida constar dessa lista de dados, ela é descartada. Caso contrário, o nó adiciona a mensagem na lista e a reenvia para todos os seus vizinhos que possuem um gradiente definido. Dessa forma, os dados fluem das fontes até o sorvedouro por múltiplos caminhos. Esses eventos iniciais são denominados de exploratórios, pois se destinam a configuração e descoberta de rotas. Após o sorvedouro começar a receber estes eventos exploratórios, ele deve reforçar positivamente um vizinho em particular.

Para reforçar positivamente um caminho, o sorvedouro envia novamente a mensagem de interesse original, porém com uma taxa de envio de dados maior. Quando o nó vizinho recebe este interesse, ele percebe que o mesmo já consta da lista de interesses, mas com uma taxa inferior. Se esta nova taxa de envio é maior do que qualquer outro gradiente para aquele interesse, o nó deve reforçar positivamente, no mínimo, um de seus vizinhos. A regra de reforço positivo é localizada, isto é, cada nó decide qual vizinho terá seu gradiente reforçado. Através dessas interações locais é estabelecido um caminho para transmissão de dados do nó fonte até o sorvedouro. Por exemplo, os nós poderiam reforçar o gradiente do primeiro nó vizinho que respondesse a um interesse propagado na rede. Nesse caso, a regra de reforço positivo garante a seleção empírica da rota de menor retardo, como ilustrado na figura 3.3(b). Essa regra é extremamente reativa a mudanças na qualidade do caminho, ou seja, sempre que ocorrer um novo caminho com retardo inferior ao estabelecido, o sorvedouro irá reforçá-lo. Em alguns casos isso poderá representar perda de recursos. Outro critério mais sofisticado inclui a escolha do nó vizinho que teria maior probabilidade de promover uma agregação de dados.

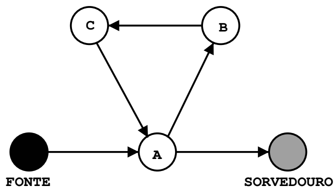
As regras descritas anteriormente também são válidas para múltiplas fontes. Na figura 3.3(c), por exemplo, os dados das fontes A e B alcançam o sorvedouro via os vizinhos C e D. No entanto, se a rota através do nó C possui consistentemente um retardo menor, somente este caminho será reforçado. Da mesma forma, se dois sorvedouros expressam interesses idênticos, a propagação dos interesses, os gradientes e as regras de reforço funcionam corretamente e os caminhos de melhor qualidade entre os sorvedouros e a fonte serão reforçados, como ilustrado na figura 3.3(d).

Os caminhos estabelecidos podem falhar ou ter sua qualidade degradada, o que exige que o nó realize um reparo local. Considerando a figura 3.3(e), quando o nó C percebe uma baixa da qualidade dos eventos enviados pela fonte, ele pode aplicar a regra de reforço para descobrir um novo caminho, ou ainda, reforçar negativamente o nó fonte. Esta degradação pode ser percebida pela diminuição da taxa de envio dos dados ou relatos de nós vizinhos indicando uma nova posição do evento monitorado. As principais causas para falha ou degradação dos caminhos são a diminuição dos níveis de energia, o mau funcionamento dos componentes do nó, a destruição do nó e os fatores ambientais que afetam a comunicação.

O algoritmo para o estabelecimento das rotas pode resultar em mais de um caminho reforçado, como ilustrado na figura 3.4(a). Essa situação ocorre quando o sorvedouro reforçar o nó A e, posteriormente, após receber um novo evento exploratório do nó B com um retardo menor, reforçar este caminho. Se a rota através do nó B for consistentemente melhor (isto é, tiver um atraso inferior ao de A), o sorvedouro precisa reforçar negativamente o caminho através do nó A para evitar o desperdício de energia. Existem vários mecanismos para aplicar o reforço negativo. Por exemplo, pode-se definir um tempo de vida para todos os gradientes, o que exige que o sorvedouro reforce periodicamente os melhores caminhos, deixando expirar os gradientes com pior qualidade. Outra abordagem, é degradar explicitamente os caminhos com desempenho inferior, enviando uma mensagem de reforço negativo. O reforço negativo também pode ser usado para remover laços como ilustrado na figura 3.4(b), uma vez que estes caminhos terão um atraso maior.



(a) Múltiplos caminhos.



(b) Caminhos com laços.

Figura 3.4: Reforço negativo.

3.2 Agregação de Dados

O conceito de fusão de dados é análogo ao processo cognitivo usado pelos seres humanos para integrar os dados oriundos dos sensores do corpo (olhos, ouvidos, nariz, língua e dedos) com um conhecimento pré-existente para inferir, de forma mais precisa, sobre os eventos que ocorrem ao seu redor. Embora essa metáfora ajude no entendimento da idéia da fusão de dados, a falta de um padrão formal tem dificultado os avanços na área. O modelo de fusão mais difundido foi criado em 1987, pelo *Joint Directors of Laboratories Data Fusion Subpanel* (JDL DFS) com fomento do Departamento de Defesa Americano. Esse modelo foi aprimorado por Waltz and Llinas [1] e, desde então, novas revisões vêm sendo propostas como, por exemplo, a apresentada por Steinberg *et al.* [72] que sintetiza fusão de dados como o processo de combinar dados para refinar estimativas e previsões.

O judicioso consumo de energia imposto às RSSFs provocou uma sutil modificação na semântica da fusão de dados empregada em sensores tradicionais (como os radares e os imageadores espectrais) passando a preponderar o termo agregação de dados. Kalpakis *et al.* [73] explicam que o conceito básico da agregação em RSSFs é combinar dados oriundos de diferentes nós sensores para eliminar as transmissões redundantes e permitir uma visão multidimensional do ambiente a ser monitorado.

A seguir serão discutidas as principais questões sobre agregação de dados com relação aos seus impactos em RSSFs.

3.2.1 Impactos da Agregação de Dados na Rede

A economia de energia proporcionada pela agregação de dados tem sido explorada e confirmada através de modelos matemáticos e experimentos envolvendo RSSFs. Entretanto, a agregação envolve uma relação de custo e benefício, a qual não deve ser negligenciada, pois pode afetar o desempenho da rede. Diversas pesquisas [5, 6, 3, 4] foram realizadas para avaliar os impactos da agregação de dados em RSSFs.

Os mecanismos de agregação de dados baseados no procedimento tradicional de armazenar e reencaminhar mensagens podem ocasionar um retardo maior na rede [5]. Esse atraso pode ocorrer porque os dados das fontes mais próximas têm que esperar aqueles mais afastados em algum nó intermediário (agregador). No pior caso, o retardo devido à agregação será proporcional ao número de saltos entre o sorvedouro e a fonte mais afastada. Caso seja considerado, ainda, que a agregação poderá ocorrer em diversos nós entre a fonte e o sorvedouro, um esquema adicional de sincronização deve coordenar quando começa e termina o processo de agregação em cada nó [6]. Outros impactos que a agregação pode ocasionar na rede de sensores diz respeito ao grau de robustez da rede, analisado em [3], e ao gasto em energia para o estabelecimento e manutenção da árvore de agregação, discutido em [4].

Os aspectos supracitados são considerados no mecanismo de agregação de dados discutido no capítulo 4.

3.2.2 Técnicas Paramétricas de Agregação de Dados

A complexidade do estudo da agregação de dados é motivada pela dependência das RSSFs de suas aplicações e a diversidade de técnicas existentes para combinar dados. Por essa razão este estudo tem como alvo a detecção, a classificação e a identificação de eventos, o que abrange uma parte significativa das aplicações em RSSFs envolvendo múltiplos sensores.

A função mais trivial de agregação de dados é a supressão de mensagens duplicadas, a qual já se encontra implementada em algumas redes comerciais [3]. Outros métodos básicos incluem calcular o valor máximo, mínimo ou médio das mensagens com o mesmo significado [74]. Estas técnicas são adequadas para agregar dados de um mesmo tipo como, por exemplo, os que permitem medir a temperatura. A agregação de dados de tipos diferentes como, por exemplo, os que indicam a presença de fumaça com aqueles que permitem medir a umidade relativa do ar, requer métodos mais complexos. Exemplos desses métodos são as técnicas oriundas do modelo JDL DFS de fusão de dados. Estas técnicas são baseadas em métodos matemáticos e heurísticos envolvendo diversas disciplinas como a estatística, a inteligência artificial, o processamento de sinais e a teoria de decisão [1].

O modelo JDL DFS classifica as técnicas de fusão de dados para detecção, identificação e classificação de eventos em: modelos físicos como, por exemplo, o Filtro de Kalman; os modelos baseados em conhecimento como, por exemplo a Teoria de Conjuntos Difusa (*Fuzzy Set Theory*); e as técnicas paramétricas que inclui a Inferência Bayesiana [75]. Este último grupo é mais apropriado para a agregação de dados em RSSFs no nível de roteamento, pois permite o mapeamento direto entre o fenômeno que se quer observar (por exemplo, *fogo*) e as evidências que caracterizam tal fenômeno (por exemplo, *temperatura*, *unidade*, *fumaça*, *monóxido de carbono* e *dióxido de carbono*). Dessa forma, a técnica de agregação explora com simplicidade e eficiência o esquema de endereçamento baseado em atributos, utilizado pelo protocolo de roteamento centrado em dados. Nesse esquema, são disseminados na rede os interesses pelas características do fenômeno, para posteriormente agregar as respostas dos sensores visando corroborar ou refutar a hipótese sobre a ocorrência do fenômeno.

As técnicas paramétricas de fusão de dados incluem a Inferência Clássica, a Bayesiana, a de Dempster-Shafer e o Processamento de Evidências Generalizado (GEP) [76]. A Inferência Clássica busca determinar a validade de uma hipótese proposta (versus uma hipótese alternativa) baseada em probabilidade empírica. A maior dificuldade dessa abordagem é a obtenção de uma função de densidade que descreva a observação do sensor, que é usada para classificar o evento de interesse. A Inferência Bayesiana resolve algumas dificuldades da Inferência Clássica como, por exemplo, permitir que múltiplas hipóteses alternativas possam ser avaliadas simultaneamente. A Inferência de Dempster-Shafer é uma generalização da anterior, mas que permite tratar a incerteza das observações dos sensores. O GEP introduz uma nova generalização da Inferência Bayesiana. Tal generalização é análoga ao método de Dempster-Shafer, no entanto, desenvolve um modo ótimo de combinar as probabilidades das proposições dos sensores. Essa otimização pode acarretar um aumento da complexidade computacional.

Na próxima seção são exploradas, com maiores detalhes, as técnicas paramétricas da Inferência Bayesiana e de Dempster-Shafer. Embora estas técnicas tenham sido foco de diversas pesquisas envolvendo o modelo JDL DFS, elas têm permanecido restritas às citações de seu grande potencial na literatura de RSSFs. A Inferência Bayesiana e de Dempster-Shafer serão empregadas no mecanismo de agregação discutido no próximo capítulo.

Inferência Bayesiana

A Inferência Bayesiana deriva do nome do clérigo inglês Thomas Bayes, autor da fórmula matemática conhecida por Teorema de Bayes [77]. Trata-se de uma técnica estatística de agregação de dados que resolve algumas dificuldades encontradas com a metodologia de Inferência Clássica [76]. A Inferência Bayesiana permite atualizar a crença de uma hipótese em função de outra estimada anteriormente e informações adicionais observadas. Essa técnica pode ser baseada na probabilidade clássica ou subjetiva.

As probabilidades subjetivas são utilizadas comumente no raciocínio humano. É comum aos seres humanos realizarem julgamento sobre determinado assunto, atribuindo uma crença para a ocorrência de eventos futuros. Esta crença é subjetiva, não repetitiva e varia de pessoa para pessoa. E ainda, a forma como é apresentada a questão (isto é positiva ou negativamente) também pode resultar em diferentes crenças. Embora a probabilidade subjetiva não possua o rigor matemático ou interpretação física encontrada na probabilidade empírica, se empregada judiciosamente, pode ser muito útil no processo de inferência da agregação de dados [75].

A Inferência Bayesiana é aplicada para encontrar a probabilidade de uma determinada hipótese ser verdadeira, considerando a existência de uma ou mais evidências. Dessa forma, supondo que H_1, H_2, ..., H_i representem hipóteses mutuamente exclusivas e exaustivas que expliquem um evento (observação) E que acabou de ocorrer, então a equação 3.1 permite calcular a crença para cada hipótese H_i formulada.

$$p(H_i|E) = \frac{p(E|H_i)p(H_i)}{\sum_i[p(E|H_i)p(H_i)]} \tag{3.1}$$

onde,

$p(H_i|E)$ = probabilidade a *posteriori* da hipótese H_i ser verdadeira dada a evidência E;
$p(E|H_i)$ = probabilidade de observar a evidência E dado que H_i é verdadeira;
$p(H_i)$ = probabilidade a *priori* da hipótese H_i ser verdadeira;
$\sum p(H_i) = 1$; e
$\sum_i[p(E|H_i)p(H_i)]$ = probabilidade de observar a evidência E dado que a hipótese H_i é verdadeira (para todas os valores de i).

Hall [75] exemplifica em seu livro uma aplicação simples de como a Inferência Bayesiana pode ser utilizada. No exemplo, a técnica é usada para calcular a probabilidade de um paciente ter realmente uma doença que obteve resultado positivo no exame laboratorial. Sabendo-se que o exame possui uma acurácia de 98% (isto é, se um paciente tem a doença, o exame indicará positivo em 98% das vezes e o falso alarme será de 2%) e que a taxa de ocorrência dessa doença é de 1 em cada 1000 pessoas. Assim, aplicando a equação 3.1, a probabilidade do paciente ter a doença é de cerca de 5%.

$$P(H) = \frac{(0,98)(0,001)}{(0,98)(0,001) + (0,02)(0,999)} = aproximadamente\ 0,05$$

Processo de Agregação de Bayes

A figura 3.5 ilustra o processo de Inferência Bayesiana aplicada em um cenário onde múltiplos sensores observam dados paramétricos (como por exemplo, assinaturas de radar, térmicas, acústicas, visuais e magnéticas) a respeito de um objeto (ou evento) cuja identidade seja desconhecida. Cada sensor declara sua hipótese acerca do evento baseado nas observações realizadas e no seu algoritmo de classificação específico. As características de desempenho anteriores (isto é, o conhecimento a *priori*) do algoritmo de classificação de cada sensor (teóricos ou experimentais) permitem estimar P(H|O_j), ou seja, a probabilidade do sensor declarar que o objeto sensoriado é de um tipo, quando na verdade é de outro diferente. Estas declarações são combinadas usando a equação 3.1 para produzir a probabilidade conjunta para cada entidade O_j baseada nas declarações dos múltiplos sensores. Portanto, o objeto j observado (dentre um conjunto de M objetos) dado as hipóteses H_1 do sensor 1, H_2 do sensor 2, etc., será o maior valor obtido para:

$$P(O_j|H_1 \cap H_2 \cap H_3 \cap ... \cap H_N),\ \text{onde j = 1, 2, ..., M}$$

As principais dificuldades para empregar a Inferência Bayesiana são a habilidade em obter P(E|H_i), ou seja, a probabilidade de observar uma evidência E assumindo uma hipótese H_i como verdadeira para cada sensor, e P(H_i) que é a probabilidade a *priori* da hipótese H_i ser verdadeira. Quando esta última não existe, o princípio da indiferença é usado, no qual o P(H_i) é igual para todos os valores de i.

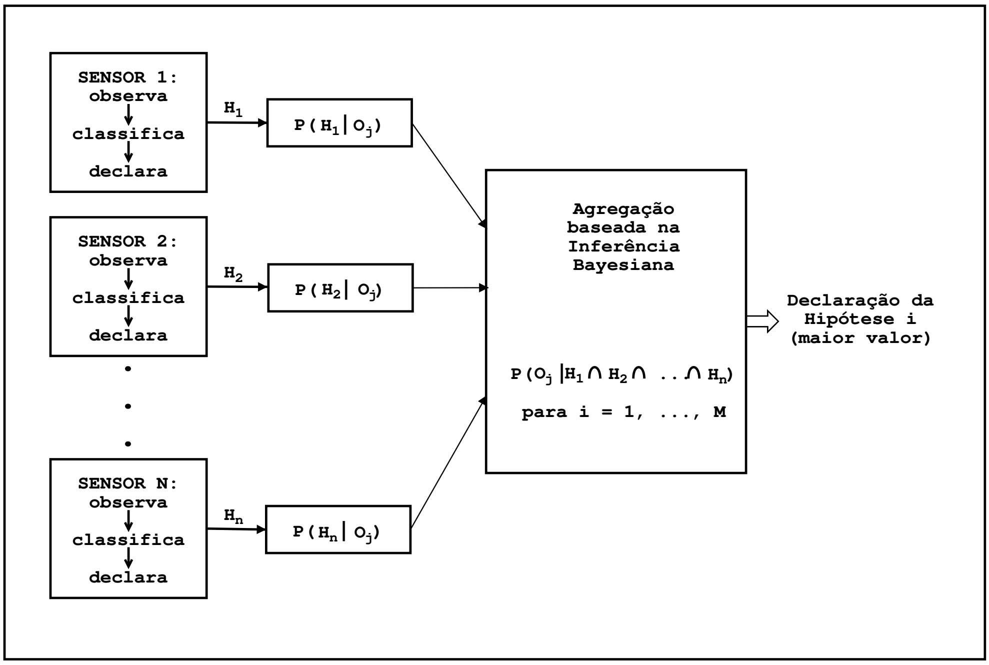


Figura 3.5: Processo de agregação baseado na Inferência Bayesiana (Adaptado de Waltz *et al.* [1]).

Inferência de Dempster-Shafer

A Inferência de Dempster-Shafer (DS) é uma generalização da Teoria de Bayes, que faz uso de intervalos de probabilidade e de incerteza para determinar a crença nas hipóteses baseadas nas observações dos sensores. O método DS combina a informação disponível independente do grau de certeza fornecida pelo sensor. Toda informação proveniente dos sensores (denominada proposições) é agregada com base na regra de Dempster [78] para determinar a intersecção ou conjunção das probabilidades associadas [76, 75].

O objetivo final da metodologia é calcular a crença nas hipóteses formuladas sobre o fenômeno observado. O conjunto das hipóteses primitivas (ou proposições elementares) que explicam um evento é chamado de quadro de discernimento (*frame of discernment*) e é denotado por Θ. Por exemplo, supondo que os eventos e_1, e_2 e e_3 representem o conjunto de todas as proposições elementares possíveis a respeito de um fenômeno observado, então $\Theta = \{e_1 , e_2 , e_3\}$. O método assume para qualquer quadro de discernimento Θ, que este é exaustivo (ou seja, contém toda hipótese primitiva possível) e que as hipóteses primitivas são mutuamente exclusivas. Assim, o conjunto Θ formado por n hipóteses primitivas dá origem à 2^n hipóteses possíveis. No exemplo anterior, com $n = 3$, existem oito proposições possíveis, denotadas por $2^\Theta = \{e_1,\ e_2,\ e_3,\ e_1 \text{ U } e_2,\ e_1 \text{ U } e_3,\ e_2 \text{ U } e_3,\ e_1 \text{ U } e_2 \text{ U } e_3,\ \emptyset\}$.

A crença em cada hipótese corresponde à probabilidade básica ou de massa (que é uma generalização da função de densidade de probabilidade) associada pelo sensor. A probabilidade básica, denotada por m, atribui um valor a cada subconjunto de Θ, onde a soma total deve ser um. Supondo que p seja uma proposição possível, o valor $m(p)$ corresponde à crença total que é atribuída exclusivamente à p. A probabilidade básica não atribuída a p, ou seja, 1 - $m(p)$, representa a incerteza de p, e é denotada por $m(\Theta)$.

A teoria de Dempster-Shafer considera que o impacto de uma evidência observada em uma proposição específica p depende de sua crença $m(p)$ e de sua negação $m(\overline{p})$. Estes dois valores são usados, respectivamente, para calcular o grau de suporte $S(p)$ e de plausibilidade $Pl(p)$ da proposição. O suporte de uma proposição p define a soma de todas as probabilidades básicas atribuídas diretamente pelo sensor para as evidências E_k que confirmam a proposição (equação 3.2), enquanto que a plausibilidade de uma proposição p fornece a soma de toda probabilidade que não é atribuída a sua negação, ou seja, a medida máxima de suporte que pode ser atribuída a proposição (equação 3.3). A faixa de valores expressa por $[S(p), Pl(p)]$ define o intervalo de confiança no qual é possível acreditar em p, sem incorrer em erros graves de suposição. Esse intervalo é tão mais amplo quanto mais incerteza houver na crença em p.

$$S(p) = \sum_{E_k \subseteq p} m(E_k) \tag{3.2}$$

$$Pl(p) = 1 - \sum_{E_k \cap p = \emptyset} m(E_k) \tag{3.3}$$

Para cada possível proposição p, a Inferência de Dempster-Shafer permite combinar as probabilidades básicas m_i e m_j das observações dos sensores S_i e S_j aplicando a equação 3.4. Esta regra pode ser generalizada para vários sensores, considerando m_j como o resultado da combinação anterior das probabilidades básicas dos sensores S_u e S_v.

$$(m_i \oplus m_j)(p) = \frac{\sum_{E_k \cap E_l = p} m_i(E_k) m_j(E_l)}{1 - \sum_{E_k \cap E_l = \emptyset} m_i(E_k) m_j(E_l)} \tag{3.4}$$

O problema de diagnóstico médico exemplifica uma aplicação da Inferência de Dempster-Shafer. Por exemplo, considerando que o quadro de discernimento Θ = {A, B, C, D} corresponda a todos os possíveis diagnósticos (hipóteses) para os sintomas (evidências) observados em um paciente, se um determinado sintoma m_1 indicar {A, B} com uma crença de $0,8$ e um outro m_2 indicar {B, C e D} com certeza de $0,9$, então a equação 3.4 pode ser utilizada para calcular as crenças nas hipóteses com base nas evidências m_1 e m_2, como ilustrado na tabela 3.1.

	$m_2(\{B,C,D\}) = 0,9$	$m_2(\Theta) = 0,1$
m_1({A,B}=0,8	$(m_1 \oplus m_2)(\{B\}) = 0,72$	$(m_1 \oplus m_2)(\{A,B\}) = 0,08$
$m_1(\Theta)$=0,2	$(m_1 \oplus m_2)(\{B,C,D\}) = 0,18$	$(m_1 \oplus m_2)(\{\Theta\}) = 0,02$

Tabela 3.1: Cálculo de $m_1 \oplus m_2$.

Assim, os possíveis diagnósticos seriam {B}, {B,C,D} e {A,B}, com crença de 72, 18 e 8%, respectivamente.

Processo de Agregação de Dempster-Shafer

O processo de agregação de dados baseado na Inferência de Dempster-Shafer é ilustrado na figura 3.6. Análogo ao processo de Bayes, os sensores coletam individualmente os dados paramétricos a respeito de um objeto (ou evento) de identidade desconhecida. Em seguida, com base nessas observações realizadas e no seu algoritmo de classificação específico, os sensores declaram sua hipótese a respeito do evento. Cada sensor k (onde $k = 1, 2, ..., N$) associa a declaração do tipo do evento e_i (onde $i = 1, 2, ..., n$) com a probabilidade básica m_k (e_i). A probabilidade básica

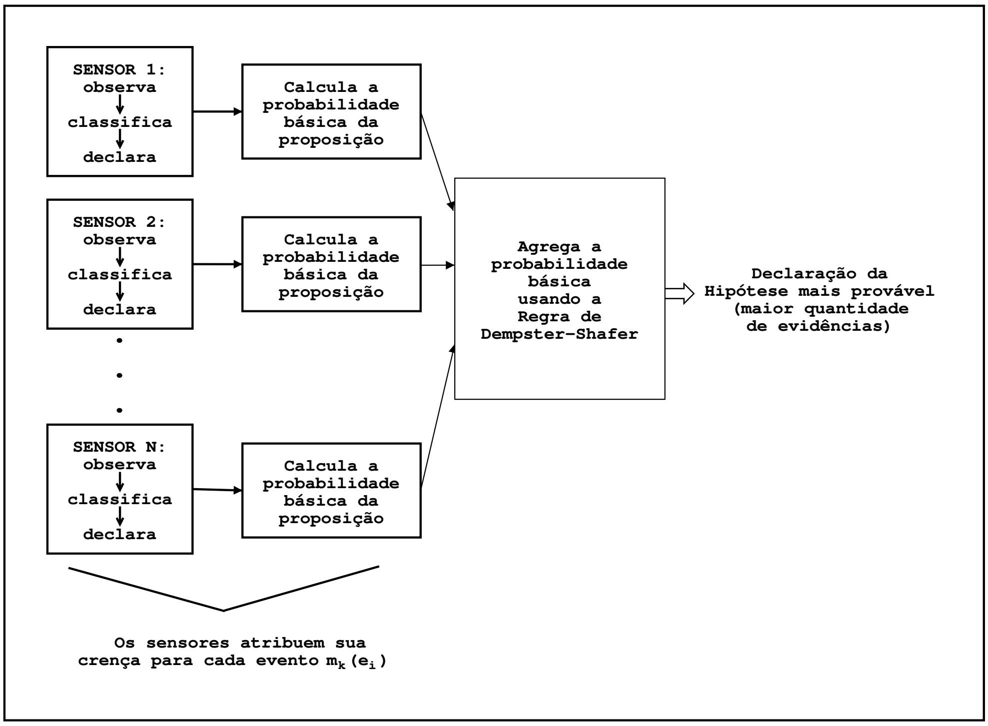


Figura 3.6: Processo de agregação baseado na Inferência de Dempster-Shafer (Adaptado de Waltz *et al.* [1]).

expressa a crença na declaração da hipótese formulada pelo sensor. Uma probabilidade básica próxima de 1 significa uma maior certeza a respeito do evento ou menos incerteza envolvida. A probabilidade básica de cada sensor é combinada usando a equação 3.4. O processo de agregação termina com a seleção da hipótese mais provável, ou seja, aquela que foi favorecida pela maior quantidade de evidências. O método permite ainda que, além dos dados coletados em tempo real, outras informações relativas ao fenômeno possam ser usadas no processo para melhorar a decisão sobre a ocorrência de uma hipótese.

3.3 Conclusão Parcial

Neste capítulo foram abordados os principais conceitos relacionados ao protocolo de roteamento em RSSFs e da agregação de dados. Este estudo ressalta a importância de combinar dados em RSSFs durante o roteamento para atingir os seguintes objetivos: economizar energia, por meio da redução da comunicação entre os nós; e ampliar o entendimento da situação monitorada, explorando informações oriundas de sensores de tipos diferentes. Tal estudo apresentou, ainda, as técnicas paramétricas de fusão de dados como métodos efetivos para alcançar tais objetivos.

O próximo capítulo mostra, em detalhes, o mecanismo de agregação de dados baseado na Inferência de Bayes e de Dempster-Shafer proposto nesta dissertação.

Capítulo 4

Mecanismo de Agregação de Dados Proposto

A agregação de dados em RSSFs envolve três diferentes aspectos, que são *onde*, *como* e *quando* agregar [2]. O primeiro aspecto é determinado pelo protocolo de roteamento, que é empregado para o estabelecimento e manutenção da árvore de agregação. Os outros dois aspectos são resolvidos pelo mecanismo de agregação que define qual técnica será utilizada e o momento em que ocorre a combinação das informações oriundas das diversas fontes.

Este capítulo apresenta o mecanismo de agregação proposto nesta dissertação. A seção 4.1 aborda a formação dos pontos de agregação utilizando o protocolo de roteamento da Difusão Direcionada. A seção 4.2 apresenta, em detalhes, os aspectos que envolvem o mecanismo de agregação de dados empregando técnicas paramétricas. Finalizando, a seção 4.3 relata as conclusões do capítulo.

4.1 Ponto de Agregação

Os protocolos de roteamento que têm por objetivo a construção de uma árvore com várias fontes convergindo para o sorvedouro enfrentam alguns obstáculos. Como foi apresentado na seção 3.2, enquanto a formação de uma árvore de agregação ótima em termos de número de transmissões é, em geral, um problema NP-completo, as soluções "quase ótimas"consomem muita energia devido à necessidade de conhecimento global da topologia da rede e ao aumento da complexidade do algoritmo de geração e manutenção da árvore de agregação. Uma outra abordagem para o problema envolve o uso de protocolos de roteamento centrados em dados. A Difusão Direcionada, que foi abordada na seção 3.1.2, é um exemplo de protocolo que utiliza esse paradigma, e que possui um elevado grau de robustez, relativa simplicidade e uma arquitetura modular. Estas características permitem influenciar heuristicamente a construção da árvore N para 1 de forma a favorecer a agregação de dados, tudo isso sem qualquer conhecimento prévio da topologia da rede.

A árvore definida pelo protocolo de roteamento gera pontos de agregação, os quais correspondem a todos os nós que recebem dados oriundos de fontes diferentes. Como ilustrado na figura 4.1, o mecanismo de agregação de dados atua nesses nós em diversos estágios até que os dados cheguem ao sorvedouro.

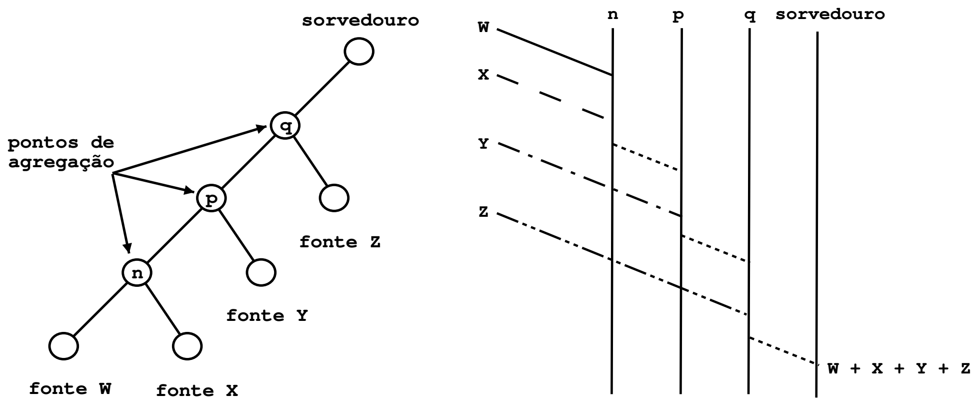


Figura 4.1: Pontos de agregação.

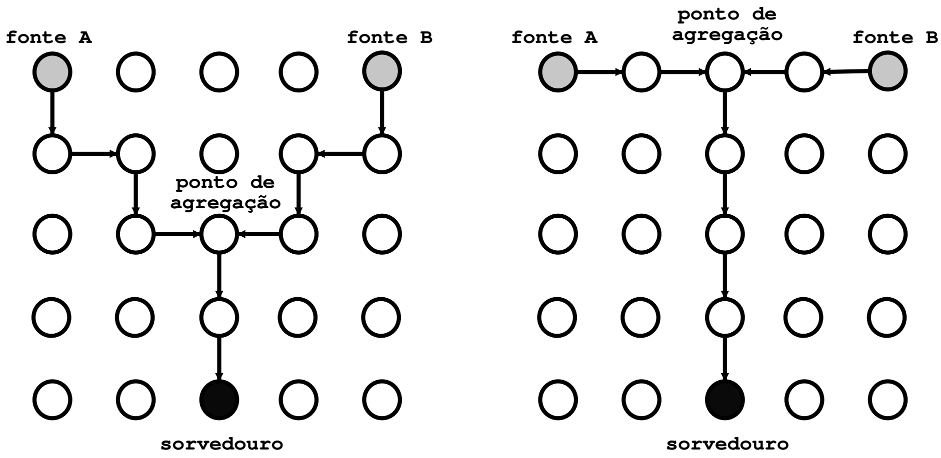
O mecanismo proposto na seção 4.2 explora a **agregação oportunista** e a **agregação privilegiada** obtida com o protocolo de roteamento da Difusão Direcionada. A **agregação oportunista** tira proveito do surgimento de pontos de agregação na formação natural das rotas de menor retardo, definidas pelo protocolo. Enquanto a **agregação privilegiada** é baseada em uma extensão da Difusão Direcionada, denominada de PAR (Roteamento Privilegiado para a Agregação) [10, 11, 79], que troca caminhos de baixa latência por outros com maior potencial de agregação. O autor desta dissertação participou da elaboração do Filtro PAR, que corresponde a uma parte dos resultados obtidos pela pesquisa em RSSFs realizada no GTA/UFRJ.

O Filtro PAR utiliza as mesmas interações localizadas características da Difusão Direcionada para estabelecer uma árvore de agregação empírica na rede. Para tal, é adicionado um módulo à implementação da Difusão Direcionada que interage com o núcleo do protocolo e modifica sua funcionalidade de descoberta de rotas. O Filtro PAR age nas mensagens exploratórias usadas para descoberta de rotas e configuração de caminhos, antes de serem processadas pelo Filtro Gradiente, que é o módulo de roteamento do protocolo. Ao invés de imediatamente encaminhar a primeira mensagem exploratória recebida, o Filtro PAR configura um temporizador a ela associado e a compara com outras até que o temporizador expire, visando selecionar fluxos com o maior número de pontos de agregação, e de forma que estes estejam próximos das fontes.

Para ilustrar as duas situações de agregação descritas, a figura 4.2 apresenta uma topologia em grade de 25 nós, com duas fontes e um sorvedouro. As rotas estabelecidas pelo protocolo da Difusão Direcionada pura seguem caminhos independentes segundo o padrão de menor retardo, variando ao longo do tempo para manter o balanceamento da energia na rede. A agregação oportunista ocorre no instante em que é formado um ponto de agregação (figura 4.2(a)). Com o PAR, as rotas também variam ao longo do tempo, no entanto deixam de ser independentes e buscam estabelecer pontos de agregação. Nesse caso, a agregação privilegiada ocorre a todo instante em que é formado um ponto de agregação (figura 4.2(b)).

4.2 Mecanismo de Agregação de Dados

O mecanismo de agregação utiliza o esquema de filtros da Difusão Direcionada, apresentada em detalhes no trabalho de Heidemann *et al.* [80], para aplicar as técnicas paramétricas propostas. Dois filtros são utilizados para a agregação, um empregando a Inferência Bayesiana e outro a Inferência de Dempster-Shafer, denominados Filtros Bayes e DS, respectivamente. A comparação entre as duas técnicas selecionadas pode ser encontrada no trabalho de David L. Hall [75].



(a) Agregação oportunista (Difusão Direcionada Pura).

(b) Agregação privilegiada (PAR).

Figura 4.2: Agregação de dados.

Os objetivos do mecanismo proposto são a economia de energia, proporcionada pela diminuição do número de transmissões na rede e, ao mesmo tempo, o refinamento das observações feitas pelos sensores, e isso com o menor impacto possível no desempenho da rede. Para atingir tais objetivos, os Filtros Bayes e DS atuam de forma descentralizada, na qual a agregação ocorre através das interações localizadas em cada ponto de agregação sem tempo de espera. O mecanismo de agregação combina a informação disponível, isto é, as observações locais e as informações obtidas dos nós vizinhos, no momento em que recebe uma mensagem. Portanto, os dados que chegam não ficam aguardando para serem agregados com outros dados oriundos de fontes que estão mais afastadas. Nesse sentido, o mecanismo difere do procedimento de armazenar e reencaminhar mensagens das técnicas tradicionais de agregação de dados. O resultado é um processamento distribuído na rede, que mantém a robustez e a escalabilidade necessária para aplicações que empregam o protocolo da Difusão Direcionada, ampliando o entendimento do fenômeno observado e reduzindo as transmissões, sem aumento significativo da latência da rede.

4.2.1 Fluxo da Mensagem de Dados

Os Filtros Bayes e DS agem nas mensagens de dados, que contém as informações de interesse coletadas pelos nós sensores, ou seja, esses filtros não interferem nas mensagens exploratórias (ou de descoberta de rotas). A figura 4.3 ilustra o fluxo da mensagem de dados no interior de um nó sensor. A mensagem chega ao núcleo da Difusão Direcionada oriunda de um nó vizinho (1) ou é gerada no próprio nó (2), quando este é uma fonte. Em seguida, segue para o filtro de agregação (3) Bayes ou DS. Depois de feita a agregação, a mensagem poderá ser descartada (4), caso seja julgada redundante, ou seguir para o núcleo (5) de onde será enviada para o Filtro Gradiente (6), que é o módulo básico da Difusão Direcionada. Este filtro é responsável pelo roteamento das mensagens para os nós vizinhos. Assim, se houver algum interesse definido, a mensagem segue para o núcleo (7) a fim de ser encaminhada para rede (8). A ordem dos filtros é determinada por um sistema de prioridades, o que implica que a prioridade da agregação deve sempre ser superior ao do roteamento.

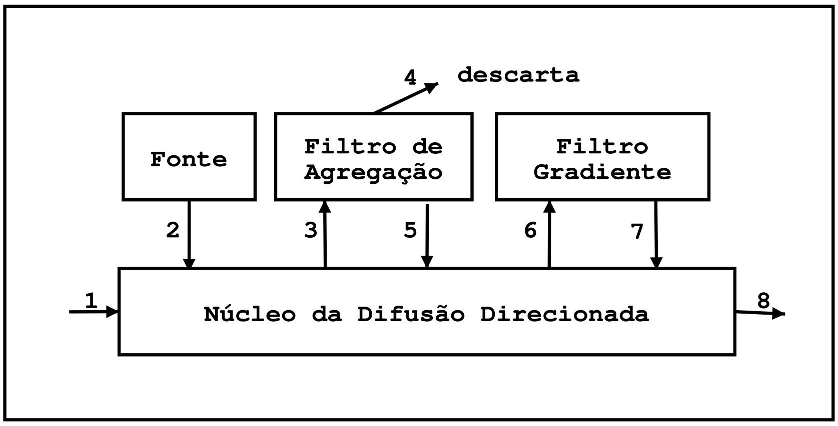


Figura 4.3: Fluxo da mensagem de dados.

4.2.2 Parâmetros do Mecanismo de Agregação

O mecanismo de agregação tem início após a definição das rotas entre os nós fonte e o sorvedouro. Dessa forma, os filtros de agregação não atuam nas mensagens exploratórias usadas para descoberta e configuração dos caminhos, tirando proveito apenas das rotas de dados com intersecção (pontos de agregação). O procedimento de agregação consiste em manter uma cópia da última evidência recebida em cada nó (do caminho fonte-sorvedouro), para ser utilizada pela função de agregação. Essa evidência pode ser usada para agregação por um período de tempo T_v, como ilustrado na figura 4.4, que determina o tempo de validade da evidência. O parâmetro T_v deverá ser inferior ao menor intervalo de transmissão (I_t) de mensagens dos nós fonte, de forma a garantir que duas mensagens de uma mesma fonte não sejam agregadas.

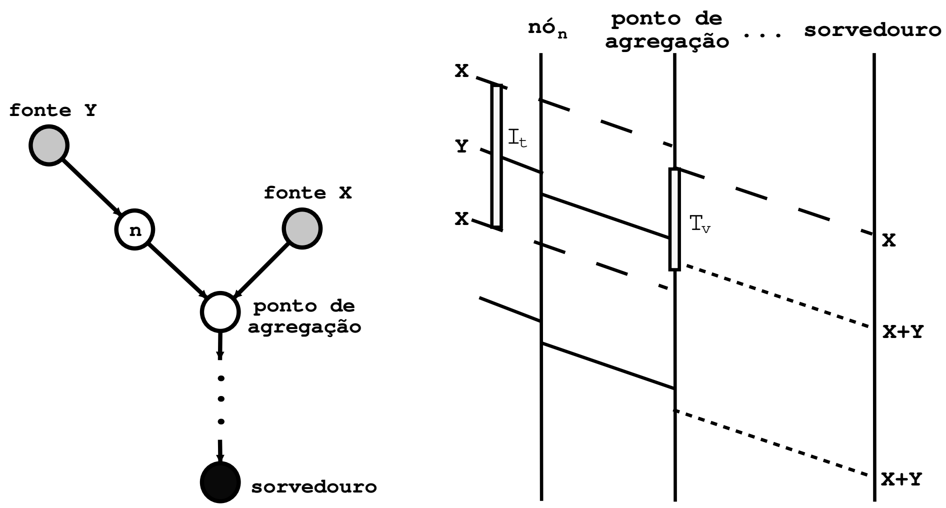


Figura 4.4: Tempo de validade.

Outro parâmetro que precisa ser ajustado é o tempo de descarte das mensagens T_d, ilustrado na figura 4.5. A idéia básica é não transmitir mensagens redundantes durante o período definido por T_d. Para tal, a última mensagem enviada é comparada com a mensagem resultante de uma nova agregação. Caso as mensagens sejam idênticas ou se nenhuma modificação significativa for introduzida, a mensagem é considerada redundante e, portanto, não será encaminhada para um nó vizinho. Por exemplo, se a agregação de uma evidência válida X com uma observação recebida Y resultar em uma mensagem $X+Y$, esta só será encaminhada para um nó vizinho se a última mensagem $X+Y$ tiver sido enviada a um intervalo superior a T_d.

No caso do Filtro DS, foi adicionado ainda o parâmetro C. Este parâmetro define a variação máxima admitida para o grau de certeza das mensagens. Dessa

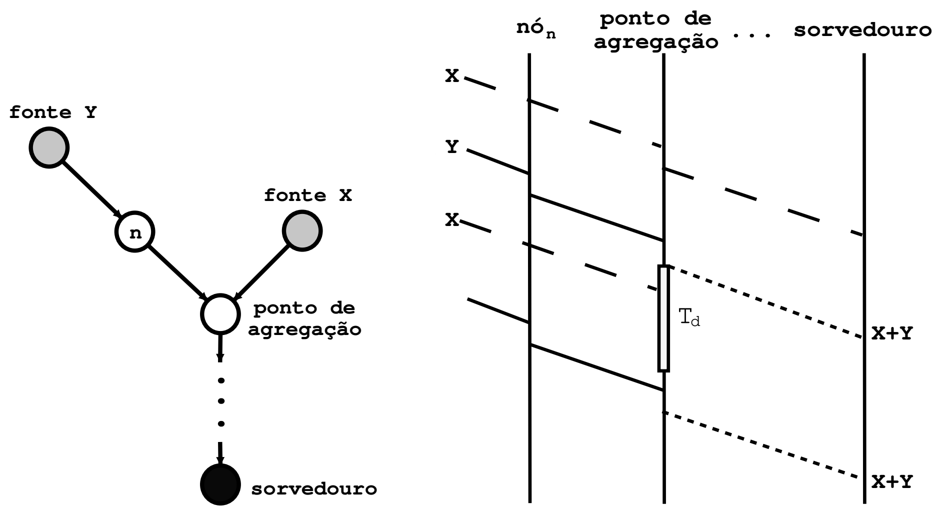


Figura 4.5: Tempo de descarte.

forma, mensagens idênticas poderão ser encaminhadas para um nó vizinho, dentro do intervalo T_d, desde que seus graus de certeza difiram de um valor superior a C. Seja o exemplo ilustrado na figura 4.6, no qual duas mensagens idênticas (P1) com diferentes graus de certeza (C_1 e C_2) são transmitidas em um intervalo inferior a T_d. Isto ocorre porque a diferença entre C_1 e C_2 é superior a C.

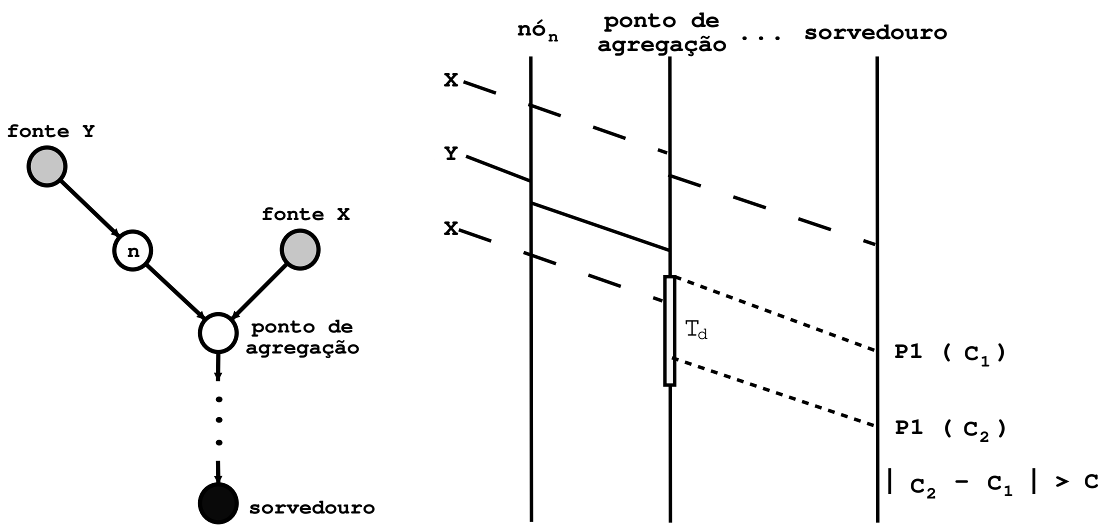


Figura 4.6: Variação máxima do grau de certeza.

4.2.3 Algoritmo dos Filtros Bayes e DS

Os nós realizam a agregação pela combinação de técnicas paramétricas de fusão de dados com o descarte de mensagens redundantes. Tais técnicas permitem um mapeamento direto entre as evidências detectadas pelos sensores e o fenômeno de interesse, o que é favorecido pelo esquema de endereçamento centrado em dados da Difusão Direcionada. Por exemplo, adotando um modelo simplificado onde o fenômeno de interesse seja "fogo" e as evidências possíveis de serem detectadas sejam "temperatura", "umidade" e "fumaça", o sorvedouro poderia difundir o interesse por "temperatura $\geq 50°C$", "umidade $\leq 30\%$" e "presença de fumaça". A técnica paramétrica seria aplicada na combinação das evidências observadas, durante o roteamento dos dados, para suportar ou refutar a crença em "fogo" na área de sensoriamento. As técnicas de agregação foram implementadas nos Filtros Bayes (com a Inferência Bayesiana) e DS (com a Inferência de Dempster-Shafer).

As figuras 4.7, 4.8 e 4.9 mostram, resumidamente, os procedimentos utilizados pelos filtros de agregação que foram implementados no simulador *ns-2* [9]. A figura 4.7 apresenta os procedimentos básicos realizados por ambos os filtros. O procedimento $configuraFiltro()$ define a prioridade do filtro e seus atributos, os quais serão usados para determinar que tipo de mensagens serão encaminhadas para a agregação. Este procedimento é realizado na fase inicial de configuração do filtro, enquanto os demais são executados sempre que uma mensagem do tipo "dado"chega ao nó.

O procedimento $validaCache()$, ilustrado na figura 4.8, consiste em verificar quais evidências do *cache* são válidas para a agregação. Isso é feito comparando a diferença entre a hora atual e o horário de chegada das evidências com o tempo de validade configurado (ou seja, o parâmetro T_v discutido na seção 4.2.2).

Em seguida, o procedimento $leMensagem()$ faz a leitura de todos os atributos da mensagem. É por meio desses atributos que o nó verifica quais evidências estão sendo reportadas e se a mensagem já sofreu alguma agregação. Após esse procedimento, as evidências válidas que estão no *cache* e as recém recebidas são processadas segundo

```
filtroAgregacao()
{
  configuraFiltro();
  validaCache();
  leMensagem(); /* recupera todos os atributos da mensagem */
  processaMensagem();
}
```

Figura 4.7: Procedimentos básicos do filtro de agregação.

a técnica de agregação implementada pelo filtro (Bayes ou DS), gerando uma nova mensagem que poderá ser roteada para a rede.

```
validaCache() \* para todas as evidências do cache */
{
  se (horaAtual - horaChegadaEvidencia) > tempoValidade
    apagaEvidencia(); /* elimina a evidência do cache */
}
```

Figura 4.8: Procedimento de validação do cache.

A figura 4.9 mostra o procedimento $processaMensagem()$ para o caso da agregação com o Filtro DS. Inicialmente, é calculado o grau de certeza de cada proposição (através da equação 3.4) disponível no nó (ou seja, as declarações dos nós sensores vizinhos e as armazenadas no *cache*). Após o *cache* ser atualizado com as novas evidências recebidas e a hora de chegada, é obtida a variação de confiabilidade de cada proposição pela diferença entre a confiança calculada e a última enviada pelo nó. Se esta variação for inferior à máxima admitida e a diferença entre a hora atual e a hora de envio da última proposição for menor que o tempo de descarte, a mensagem é descartada. Caso contrário, uma nova mensagem é gerada contendo todas as proposições e seus respectivos graus de confiança calculados, para ser encaminhada para o Filtro Gradiente. E nesse caso, a hora de envio e as proposições enviadas são atualizadas. No Filtro Bayes, o procedimento é semelhante, sendo que a agregação é baseada em hipóteses que são combinadas utilizando a equação 3.1.

```
processaMensagem() /* para Inferência de Dempster-Shafer */
{
  calculaConfiabilidadeProposicao();
  atualizaCache();
  calculaVariacaoConfiabilidade();
  se (horaAtual - horaUltimaProposicaoEnviada) < tempoDescarte
    se variacaoConfiabilidade < maximaVariacaoConfiabilidade
      descartaMensagem();
  senão
    gerar mensagemAgregada();
    enviar mensagemAgregada();
    horaUltimaProposicaoEnviada = horaAtual;
    ultimaProposicaoEnviada = proposicaoCalculada;
}
```

Figura 4.9: Procedimento para agregação das mensagens.

4.2.4 Exemplos de Aplicação

Esta subseção apresenta dois exemplos de aplicação do mecanismo de agregação proposto para RSSFs. No primeiro exemplo é utilizado a Inferência Bayesiana e no segundo a de Dempster-Shafer.

Aplicação com a Inferência Bayesiana

A aplicação consiste em monitorar as condições ambientais de um determinado local para realização de uma operação militar ou de resgate. Os sensores são lançados na área para verificar a existência de evidências (tais como, temperatura do ar e do solo, umidade relativa do ar, pressão atmosférica, direção e velocidade do vento, radiação solar, fumaça) que corroborem ou refutem as hipóteses que explicam um fenômeno de interesse (como por exemplo, incêndio, área contaminada, risco de acidente aéreo). Assim, é possível criar um modelo no qual as probabilidades a *priori* das hipóteses são estabelecidas, bem como, as de certas evidências serem observadas, dada a confirmação de uma hipótese, e então, aplicar a Inferência Bayesiana para agregar as evidências detectadas pelos nós sensores, computando a crença nas hipóteses formuladas.

Adotando um cenário genérico em que $\{H_1, H_2, H_3$ e $H_4\}$ representam o conjunto de todas as possíveis hipóteses e $\{E_1, E_2, E_3$ e $E_4\}$ correspondem às evidências que

i	1	2	3	4
$P(H_i)$	0,1	0,4	0,3	0,2
$P(E_1\|H_i)$	0,7	0,9	0,4	0,2
$P(E_2\|H_i)$	0,8	0,2	0,0	0,4
$P(E_3\|H_i)$	0,1	0,0	0,8	0,2
$P(E_4\|H_i)$	0,0	0,2	0,4	0,9

Tabela 4.1: Valores das probabilidades a *priori* e condicionais.

podem ser detectadas pelos sensores, a tabela 4.1 apresenta os valores conhecidos para as crenças a *priori* de cada hipótese e o grau com que cada evidência suporta tais hipóteses.

Após o lançamento dos sensores, o sorvedouro difunde o interesse pelas evidências e aguarda uma resposta. Se o nó fonte S_1 identifica uma evidência E_1 (e não dispõe de nenhuma outra informação), ele atualiza as crenças nas hipóteses e gera uma mensagem com esses valores, que são difundidos na rede. A tabela 4.2 mostra o resultado desse cálculo, onde se observa que se a evidência suporta a hipótese, sua crença aumenta, caso contrário, diminui. Assim, com a observação de E_1, as crenças nas hipóteses H_1 e H_2 aumentaram (de 0,1 para 0,12 e de 0,4 para 0,61, respectivamente), enquanto H_3 e H_4 diminuíram (0,3 para 0,2 e 0,2 para 0,07, respectivamente).

Quando outro nó sensor S_2 identifica uma evidência E_2, procede de maneira idêntica, atualizando a crença nas hipóteses e propagando a informação. No entanto, se existir um ponto de agregação na rede, ou seja, um nó que receba as mensagens oriundas de S_1 e S_2, a crença nas hipóteses são refeitas, como mostrado na tabela 4.3, e a mensagem agregada é roteada para o nó vizinho. Nesse caso, a agregação das evidências E_1 e E_2 faz a crença em H_3 ser zero (ou seja, pode ser desprezada), enquanto reduz a crença em H_2 e aumenta de H_1 e H_4 . As mensagens agregadas que resultam em hipóteses com a mesma crença, em um intervalo de tempo definido (T_d), são descartadas.

A crença nas hipóteses continua a ser calculada a medida que novas evidências

i	$P(H_i\|E_1)$
1	$\frac{0,7*0,1}{0,7*0,1+0,9*0,4+0,4*0,3+0,2*0,2} = 0,12$
2	$\frac{0,9*0,4}{0,7*0,1+0,9*0,4+0,4*0,3+0,2*0,2} = 0,61$
3	$\frac{0,3*0,4}{0,7*0,1+0,9*0,4+0,4*0,3+0,2*0,2} = 0,20$
4	$\frac{0,2*0,2}{0,7*0,1+0,9*0,4+0,4*0,3+0,2*0,2} = 0,07$

Tabela 4.2: Probabilidade a *posteriori* das hipóteses (dado a evidência E_1).

i	$P(H_i\|E_1E_2)$
1	$\frac{0,7*0,8*0,1}{0,7*0,8*0,1+0,9*0,2*0,4+0,0+0,2*0,4*0,2} = 0,39$
2	$\frac{0,9*0,2*0,4}{0,7*0,8*0,1+0,9*0,2*0,4+0,0+0,2*0,4*0,2} = 0,50$
3	$\frac{0,4*0,0*0,3}{0,7*0,8*0,1+0,9*0,2*0,4+0,0+0,2*0,4*0,2} = 0,0$
4	$\frac{0,2*0,4*0,2}{0,7*0,8*0,1+0,9*0,2*0,4+0,0+0,2*0,4*0,2} = 0,11$

Tabela 4.3: Probabilidade a *posteriori* das hipóteses (dadas as evidências E_1 e E_2).

i	$P(H_i \mid E_1 E_2 E_3)$
1	$\frac{0,7*0,8*0,1*0,1}{0,7*0,8*0,1*0,1+0,0+0,0+0,2*0,4*0,2*0,2} = 0,64$
2	$\frac{0,9*0,2*0,0*0,4}{0,7*0,8*0,1*0,1+0,0+0,0+0,2*0,4*0,2*0,2} = 0,0$
3	$\frac{0,4*0,0*0,8*0,3}{0,7*0,8*0,1*0,1+0,0+0,0+0,2*0,4*0,2*0,2} = 0,0$
4	$\frac{0,2*0,4*0,2*0,2}{0,7*0,8*0,1*0,1+0,0+0,0+0,2*0,4*0,2*0,2} = 0,36$

Tabela 4.4: Probabilidade a *posteriori* das hipóteses (dado as evidências E_1, E_2 e E_3).

são observadas. Assim, se um nó sensor possui em seu *cache* a informação sobre as observações E_1 e E_2 e recebe uma mensagem com a evidência E_3, torna-se um ponto de agregação e atualiza as crenças nas hipóteses com os valores apresentados na tabela 4.4. Com a agregação das observações E_1, E_2 e E_3, as hipóteses H_2 e H_3 são refutadas, enquanto a crença em H_1 e H_4 aumentam. De forma análoga ao caso anterior, uma nova mensagem, com os valores das crenças atualizadas, é gerada e encaminhada para a rede, descartando as posteriores que obtiverem resultados semelhantes.

Aplicação com a Inferência de Dempster-Shafer

O cenário da aplicação empregando a Inferência de Dempster-Shafer considera a monitoração de uma área de conflito, onde existem alvos amigos e inimigos. O objetivo da rede de sensores é identificar os alvos na região. O modelo a ser explorado conta com sensores capazes de detectar seis possíveis alvos, como discriminado a seguir:

$$a_1 = \text{alvo amigo tipo 1} \qquad a_4 = \text{alvo inimigo tipo 1}$$
$$a_2 = \text{alvo amigo tipo 2} \qquad a_5 = \text{alvo inimigo tipo 2}$$
$$a_3 = \text{alvo amigo tipo 3} \qquad a_6 = \text{alvo inimigo tipo 3}$$

Depois de propagado o interesse na rede, cada sensor irá reportar as proposições sobre os alvos detectados com um determinado grau de certeza, o qual será usado para combinar as observações dos sensores. Assim, se o sensor S_1 difunde na rede a proposição $m_1 = \{a_1 \cup a_4\}$ com uma crença de $0,7$ (ou uma incerteza de $0,3$ do alvo ser do tipo 1) e outro sensor S_2 propaga a proposição $m_2 = \{a_4 \cup a_5 \cup a_6\}$ com uma crença de $0,8$ (ou uma incerteza de $0,2$ do alvo ser inimigo), então o primeiro nó comum às duas rotas estabelecidas (ponto de agregação) entre as fontes (S_1 e S_2) e o sorvedouro poderão, em um intervalo de tempo definido, aplicar a Inferência de Dempster-Shafer para agregar as proposições e a partir daí, descartar mensagens redundantes. A tabela 4.5 mostra o resultado da agregação, o qual considera que o alvo identificado pode ser o $\{a_4\}$ ou $\{a_4 \cup a_5 \cup a_6\}$, ou ainda, $\{a_1 \cup a_4\}$, com confiança de 0,56, 0,24 e 0,14, respectivamente.

	$m_2(a_4 \cup a_5 \cup a_6) = 0,8$	$m_2(\Theta) = 0,2$
$m_1(a_1 \cup a_4) = 0,7$	$m(a_4) = 0,56$	$m(a_1 \cup a_4) = 0,14$
$m_1(\Theta) = 0,3$	$m(a_4 \cup a_5 \cup a_6) = 0,24$	$m(\Theta) = 0,06$

Tabela 4.5: Agregação de m_1 e m_2.

Caso o ponto de agregação receba, no mesmo intervalo, uma nova proposição $m_3 = \{a_4 \cup a_5 \cup a_6\}$ com uma crença de $0,9$, a Inferência de Dempster-Shafer é aplicada novamente, utilizando as proposições calculadas anteriormente, como mostra a tabela 4.6. Nesse caso, a mensagem agregada resultaria nas proposições $\{a_4\}$ ou $\{a_4 \cup a_5 \cup a_6\}$ ou $\{a_1 \cup a_4\}$ com a confiança de 0,69, 0,29 e 0,01, respectivamente. Esse resultado é comparado com a última mensagem enviada para decidir se as novas proposições obtidas serão descartadas ou não.

	$m_3(a_4 \cup a_5 \cup a_6) = 0,9$	$m_3(\Theta) = 0,1$
$m(a_4) = 0,56$	$m^{'}(a_4) = 0,50$	$m^{'}(a_4) = 0,06$
$m(a_1 \cup a_4) = 0,14$	$m^{'}(a_4) = 0,13$	$m^{'}(a_1 \cup a_4) = 0,01$
$m(a_4 \cup a_5 \cup a_6) = 0,24$	$m^{'}(a_4 \cup a_5 \cup a_6) = 0,22$	$m^{'}(a_4 \cup a_5 \cup a_6) = 0,02$
$m(\{\Theta\}) = 0,06$	$m^{'}(a_4 \cup a_5 \cup a_6) = 0,05$	$m^{'}(\Theta) = 0,01$

Tabela 4.6: Agregação de m_1, m_2 e m_3.

4.3 Conclusão Parcial

Este capítulo apresentou o mecanismo de agregação proposto para RSSFs em aplicações que empregam o protocolo da Difusão Direcionada. O mecanismo preserva a característica do protocolo que garante uma elevada robustez da rede. Adicionalmente, tira proveito do paradigma de comunicação centrado em dados, para facilitar a agregação empregando técnicas paramétricas, com um mínimo impacto na rede. Também foi explorada a criação de pontos de agregação de forma oportunista e privilegiada. Na parte final, foram abordados os filtros implementados no módulo da Difusão Direcionada do simulador *ns-2* e dois exemplos numéricos que ilustraram a aplicação dessas técnicas. O capítulo seguinte avalia, por meio de simulações, o mecanismo proposto.

Capítulo 5

Simulações e Resultados

Este capítulo apresenta os principais resultados obtidos de simulações, realizadas visando verificar se o mecanismo de agregação proposto atende aos seus objetivos. As avaliações preliminares feitas pelo autor deste trabalho de dissertação foram publicadas em [10]. Uma avaliação mais detalhada desse mecanismo será publicada em [11].

A próxima seção descreve o ambiente de simulação, o módulo da Difusão Direcionada implementado no simulador, os filtros e as aplicações adicionadas a este módulo, e a configuração dos cenários de simulação montados para cada filtro com o objetivo de avaliar o mecanismo proposto. A seção 5.2 apresenta os resultados obtidos com as métricas de desempenho. Por último, a seção 5.3 traz as conclusões do capítulo.

5.1 Ambiente de Simulação

O ambiente de simulação utilizado para as avaliações foi o *Network Simulator* (*ns-2*) versão 2.26 [9]. O *Network Simulator* é um simulador orientado a eventos discretos utilizado em pesquisas de redes de comunicação. Foi desenvolvido no projeto VINT (*Virtual InterNetwork Testbed*) com apoio da DARPA (*Defense Advanced Research Projects Agency*) através do LBL (*Lawrence Berkley National Laboratory*), da UCB (*University of California in Berkley*), da USC/ISI (*University of Southern California / Information Sciences Institute*) e do Xerox PARC (*Palo Alto Research Center*). Atualmente, o *ns-2* é mantido pela DARPA com o SAMAN (*Simulation Augmented by Measurement and Analysis for Networks*) e pelo NSF (*National Science Foundation*) com o CONSER (*Collaborative Simulation for Education and Research*), ambos em colaboração com outros pesquisadores de universidades e centros de pesquisa.

O *ns-2* utiliza as linguagens C++ e OTcl (*Object Tool Command Language*) para assegurar desempenho e flexibilidade ao simulador. Dessa forma, para garantir uma maior velocidade de execução dos algoritmos que processam um grande conjunto de eventos, o núcleo do simulador foi implementado em C++, enquanto a interface de comando e configuração utiliza *scripts* escritos em OTcl, que permitem realizar modificações nos parâmetros e nos cenários de simulação mais rapidamente.

5.1.1 Módulo da Difusão Direcionada no ns-2

O módulo da Difusão Direcionada no *ns-2* é baseado na implementação do grupo SCADDS da USC/ISI. Este módulo consiste de uma arquitetura de duas camadas e uma biblioteca de interface entre elas. A primeira camada é chamada de núcleo da difusão e é usada para enviar e receber os pacotes da rede. A segunda camada é a de aplicação, na qual ficam os filtros e as aplicações. A biblioteca de interface possibilita a comunicação entre as camadas por meio de duas API [81]: a *Network* e a *Filter*.

A *Network* API é baseada no paradigma subscrever/publicar dados. A interface permite realizar dois tipos de operações, com as quais os nós podem **subscrever** dados, ou seja, manifestar o interesse por determinados eventos, e **publicar** dados, isto é, informar a disponibilidade de eventos. Os dados são nomeados usando um conjunto de atributos, onde cada um possui vários campos, tais como: a *chave* (que contém a semântica do atributo), o *operador* (que descreve como comparar dois atributos), o *tipo* (que indica a natureza do atributo) e o *valor* (que define o conteúdo do atributo).

A *Filter* API é o mecanismo responsável pelo desempenho da rede. As aplicações são adicionadas à difusão como módulos de software (que são chamados de filtros) e influenciam como os dados se movem através da rede. Alguns exemplos de filtros incluem funções de roteamento e de agregação de dados. Cada filtro possui uma prioridade, que define a ordem de processamento, e uma lista de atributos, que determinam quais pacotes serão processados.

5.1.2 Filtros e Aplicações Adicionadas ao Módulo da Difusão Direcionada

Para realizar as simulações do mecanismo proposto nesta dissertação foram adicionados ao módulo de difusão, por meio da *Network e Filter* API, os filtros de agregação (Bayes e DS) e as aplicações que geram os dados na rede, como ilustrado na figura 5.1. Dessa forma, evitou-se modificações diretas no núcleo da difusão e no Filtro Gradiente (responsável pelo roteamento), preservando o elevado grau de adaptabilidade da Difusão Direcionada, que lhe permite recuperar rotas em caso de mudanças na topologia.

Os Filtros Bayes e DS foram escritos em C++, totalizando cerca de mil e quinhentas linhas de código, e implementam as classes *FusionFilter* e *DSFusionFilter*, respectivamente. Estas classes recebem as mensagens de dados e aplicam a técnica paramétrica de agregação específica do filtro. Os parâmetros dos filtros (T_v, T_d e C) podem ser modificados pelos *scripts* de simulação para ajustar o mecanismo conforme a aplicação.

Para a avaliação das técnicas de agregação, houve a necessidade de se implementar módulos de aplicação responsáveis pela geração dos pacotes de interesse e de dados, específicos para o estudo realizado. Estas aplicações foram implementadas com base no código do *PingApplication* [82], desenvolvido por Fábio Silva, da USC/ISI. O *Ping Application* possui duas classes denominadas *PingSenderApp* e *PingReceiverApp*, que usam um padrão de projeto denominado "fábrica de obje-

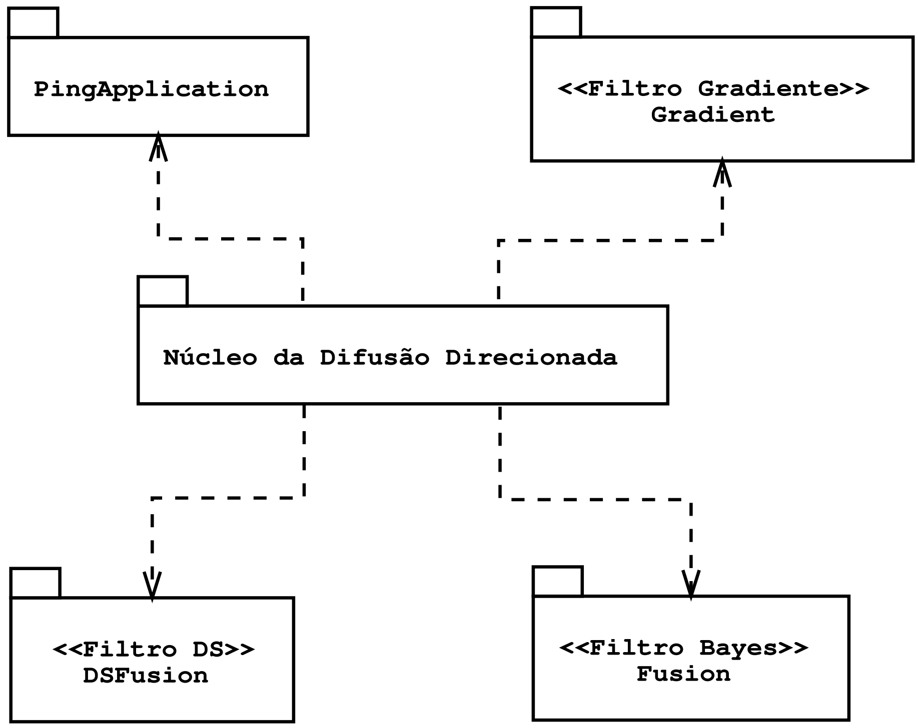


Figura 5.1: Filtros e aplicações do módulo da Difusão Direcionada.

tos" (*factory method*) definido nas APIs para criação dos atributos dos interesses e dos dados que fluem na rede. A aplicação foi adaptada para cada filtro de agregação e modificada para atender às necessidades específicas de cada cenário de simulação.

5.1.3 Configuração dos Cenários de Simulação

Os principais parâmetros utilizados para configuração dos cenários de simulação foram: para o protocolo de roteamento, o *Directed Diffusion*; para a camada de enlace, o protocolo 802.11; para o modelo de propagação rádio, o *Two-ray ground*; e a interface de rádio foi baseada no *914MHz Lucent WaveLAN DSSS* , com o alcance do rádio definido em um valor constante de 250 metros. Os cenários consideram uma rede plana e sem mobilidade. A composição da RSSF considera dois tipos de sensores diferentes como, por exemplo: para o Filtro Bayes, medição da temperatura e concentração de monóxido de carbono do ar; e para o Filtro DS, detecção da assinatura acústica e magnética de um objeto.

Os módulos de aplicação implementados para os diversos cenários de simulação, geram mensagens de dados a cada segundo e o período de simulação em todas as rodadas foi de 200 segundos. Foram definidos cenários com uma distribuição de nós regular e aleatória. Na distribuição regular, os nós formam uma topologia em forma de grade, afastados entre si de 200 metros, restringindo a comunicação com o nó vizinho em diagonal. Os cenários de simulação desenvolvem-se em um campo de sensoriamento que ocupa uma área quadrada, com um único sorvedouro (situado no canto inferior esquerdo do campo) que é responsável pela disseminação das tarefas para os nós sensores, e pelas fontes (distribuídas na rede conforme o cenário avaliado) que respondem aos interesses.

Para as simulações realizadas, foi criado um modelo de aplicação para cada filtro de agregação. O Filtro Bayes foi implementado para uma aplicação genérica, de forma a atender a inúmeros problemas de agregação. Este modelo consiste de um conjunto de quatro hipóteses mutuamente exclusivas e exaustivas que possuem sua crença aumentada ou diminuída segundo a ocorrência de um ou dois eventos possíveis de serem detectados pelos nós sensores. As probabilidades *a priori* de cada hipótese e as probabilidades de cada hipótese ser verdadeira dado que uma evidência ocorreu, foram definidas aleatoriamente. O modelo de aplicação do Filtro DS simula a classificação de quatro possíveis alvos detectados na área de sensoriamento. Cada nó fonte que detecta um alvo, transmite sua proposição com um grau de certeza que é gerado aleatoriamente a cada envio.

5.2 Avaliação do Mecanismo de Agregação

Para avaliar o impacto do mecanismo de agregação em uma RSSF com a Difusão Direcionada, foram definidas as seguintes métricas de desempenho: para a economia de energia, a quantidade de mensagens entregues no sorvedouro e o número total de transmissões (saltos) na rede; e para a latência da rede, o atraso total dos pacotes entre a fonte e o sorvedouro. As métricas são comparadas com e sem o emprego do mecanismo de agregação.

A redução da quantidade de mensagens no sorverdouro é obtida com a agregação, que realiza um processamento na rede para refinar as informações dos sensores. O número de transmissões realizadas é importante, tendo em vista que o maior consumo de energia ocorre quando um nó transmite dados. Uma redução do número de mensagens no sorvedouro com diminuição das transmissões representa um indicativo de economia de energia e, considerando que a agregação realizada pelos Filtros Bayes e DS não implica em ciclos de espera, o atraso introduzido pelo mecanismo fica restrito ao tempo de processamento do filtro, minimizando assim, o impacto na latência da rede.

Com exceção dos cenários 1 e 6, as simulações foram realizadas com o Filtro Bayes explorando a formação de pontos privilegiados para a agregação (com o Filtro PAR), enquanto o Filtro DS foi empregado nos pontos de agregação oportunistas da Difusão Direcionada pura. Os conceitos de agregação privilegiada e oportunista foram apresentados na seção 4.1.

As métricas de desempenho selecionadas foram usadas nos seguintes cenários: parâmetros do mecanismo de agregação; escalabilidade; número de nós fontes; posição relativa dos nós fontes; topologia aleatória; e latência da rede. Os dados apresentados em todos os cenários foram obtidos com dez rodadas de simulação e um intervalo de confiança de 95%. Na seção 5.2.7 é apresentado um resumo dos resultados.

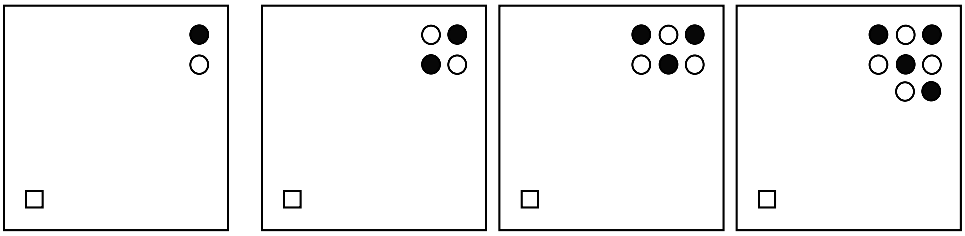
5.2.1 Cenário 1: Parâmetros do Mecanismo de Agregação

Inicialmente foi analisada a influência dos parâmetros do mecanismo de agregação no desempenho da rede. As simulações foram realizadas em uma topologia regular de 100 nós, com as fontes dispostas próximas entre si, de forma a favorecer a agregação. A figura 5.2 ilustra a distribuição dos nós. Este cenário considera a agregação privilegiada, ou seja, foram empregados o Filtros Bayes e DS junto com o PAR. A taxa média de agregação (TMA) no sorvedouro, dada pela equação 5.1, corresponde ao percentual médio de redução de mensagens obtidos com os filtros de agregação em relação a Difusão Direcionada pura.

$$TMA = 1 - \frac{M_{DD}}{M_{AG}} \tag{5.1}$$

onde,

M_{DD} = Número médio de mensagens recebidas com a Difusão Direcionada Pura; e
M_{DD} = Número médio de mensagens recebidas com filtro de agregação.



(a) 2 fontes. (b) 4 fontes. (c) 6 fontes. (d) 8 fontes.

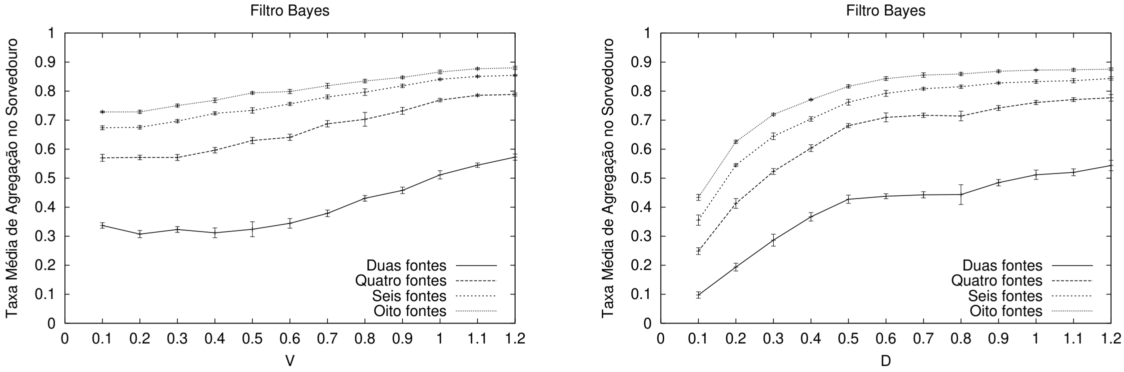
Figura 5.2: Topologia em grade com cem nós.

Nos gráficos das figuras 5.3 e 5.4 as variáveis V e D correspondem, respectivamente, a relação entre o tempo de validade da evidência e o intervalo de transmissão da fonte (T_v/I_t) e o tempo de descarte e o intervalo de transmissão da fonte (T_d/I_t), vistos em detalhes na seção 4.2.2. Na avaliação dos parâmetros do mecanismo foram usadas duas, quatro, seis e oito fontes.

Os gráficos da figura 5.3 mostram o resultado da agregação com o Filtro Bayes. O gráfico 5.3(a) indica uma tendência de aumento da taxa de agregação na medida que V cresce, sendo mantido constante o valor $D = 1,0$. Essa tendência de aumento na taxa de agregação é acompanhada por uma redução do número de transmissões, como se pode verificar na tabela 5.1. No entanto, um tempo de validade muito grande pode deixar a rede menos reativa a mudanças. No gráfico 5.3(b) o valor de $V = 1,0$ permanece fixo, enquanto D varia. Neste gráfico, observa-se que nas curvas de quatro, seis e oito fontes, um tempo de descarte baixo provoca uma elevada

Filtro Bayes								
	2 fontes		4 fontes		6 fontes		8 fontes	
V	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95
0,10	10898,00	766,71	20652,60	1383,46	28189,20	1549,71	32898,50	1213,86
0,20	11257,80	798,04	20571,00	924,24	27858,30	1332,75	33303,50	1680,24
0,30	10649,20	696,10	20765,20	1425,36	26807,30	1392,54	30284,50	1306,85
0,40	10993,80	613,31	19670,60	1177,40	24088,60	1318,35	27910,60	1323,00
0,50	10735,20	733,76	17703,80	1143,43	23536,50	1400,88	24786,20	1588,96
0,60	10281,10	527,08	17262,70	833,08	20966,60	852,87	24981,60	1346,62
0,70	10263,50	547,15	15043,00	1318,63	19345,10	859,11	22729,20	713,01
0,80	8967,30	546,58	14322,50	863,74	17320,80	1012,50	19475,50	786,00
0,90	8364,20	727,33	12737,60	603,09	15593,30	642,32	18163,70	628,20
1,00	7812,10	626,92	11111,80	663,31	13851,90	621,55	16284,00	693,39
1,10	7456,70	413,24	10012,10	234,40	13445,30	396,93	15650,30	732,72
1,20	6938,60	271,64	10046,50	590,87	13039,90	808,72	15680,90	406,90
	2 fontes		4 fontes		6 fontes		8 fontes	
D	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95
0,10	14299,80	517,32	36003,40	1328,78	52206,00	2473,75	68160,70	3610,89
0,20	12926,40	849,80	27693,30	1246,69	37886,90	1366,61	47342,30	1891,35
0,30	11304,10	616,35	22937,50	1599,83	31043,10	1385,48	37232,60	1191,91
0,40	10070,60	697,35	19226,10	1592,45	26179,60	662,84	29910,70	905,64
0,50	9123,20	354,53	15270,10	927,03	21134,60	843,33	24602,70	602,47
0,60	9028,30	361,35	15209,60	909,66	18224,80	1114,70	21129,90	1112,56
0,70	9028,20	387,25	14307,60	664,36	17169,20	846,82	19552,60	1031,79
0,80	9025,20	282,18	12975,50	679,67	17085,60	670,43	18783,40	692,14
0,90	8264,70	379,00	12053,00	675,94	15402,80	259,08	17638,90	652,90
1,00	7896,90	716,65	11535,00	665,03	14415,70	647,55	17005,80	670,59
1,10	7547,70	623,98	10908,40	441,21	13882,80	933,03	16089,00	855,82
1,20	7158,60	507,16	10849,80	702,02	12793,40	882,19	16184,80	1070,48

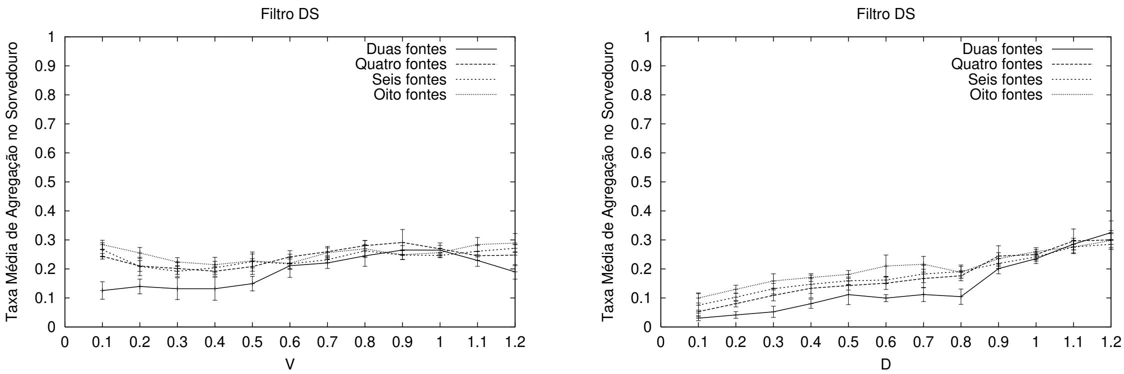
Tabela 5.1: Número de saltos no Filtro Bayes (com variação do parâmetro V e D).



(a) Tempo de validade. (b) Tempo de descarte.

Figura 5.3: Parâmetros do mecanismo de agregação com o Filtro Bayes.

taxa de agregação, o que evidencia uma concentração de mensagens com pequeno atraso entre elas. Como conseqüência, a partir de um valor para D (no caso, aproximadamente $0,5$), a taxa de agregação passa a sofrer variações menores. O mesmo comportamento pode ser observado na tabela 5.1 com relação ao número de saltos. Assim, o aumento de D favorece as métricas de desempenho. No entanto, quando este ultrapassa a unidade (ou seja, $T_d > I_t$) o intervalo de chegada das mensagens no sorvedouro pode ser elevado significativamente.



(a) Tempo de validade. (b) Tempo de descarte.

Figura 5.4: Parâmetros do mecanismo de agregação com o Filtro DS.

Os resultados da agregação com o Filtro DS são apresentados na figura 5.4. O gráfico 5.4(a) indica uma pequena diferença entre as taxas de agregação (principalmente nas curvas com mais de duas fontes), na medida em que V aumenta e D é mantido constante e igual a 1,0. Para $V = 1,0$, as quatro curvas tendem para uma taxa de agregação aproximada de 26%, sendo observado na tabela 5.2, que para esse valor de V são obtidas as menores quantidades de transmissões (sem comprometer a reatividade da rede). No gráfico 5.4(b) a taxa de agregação cresce com o aumento de D, enquanto V permanece fixo e igual a 1,0. O valor de $D = 1,0$ garante uma taxa de agregação em cerca de 25% e o menor número de saltos (sem elevar significativamente a taxa de chegada das mensagens no sorvedouro), como pode ser visto na tabela 5.2.

No caso do Filtro DS, as curvas da figura 5.5 mostram o efeito da manipulação da variação máxima admitida para o grau de certeza (C) das proposições, considerando os valores fixos para $V = 1,0$ e $D = 1,0$. Pelo gráfico, pode-se verificar que a taxa de agregação no sorvedouro cresce com o aumento de C. Isso ocorre, porque um número maior de mensagens é descartado nos pontos de agregação, devido ao relaxamento da variação do grau de certeza das proposições dos sensores. O desempenho da rede em termos de transmissões também melhora com o aumento de C, como pode ser verificado na tabela 5.3. O aumento do número de fontes não teve influência significativa na taxa de agregação, embora tenha sido expressiva no caso do número de transmissões. Por exemplo, para $C = 0,1$ a taxa de agregação foi de aproximadamente 25%, independente do número de fontes, enquanto o número de transmissões com oito fontes foi quase oito vezes maior do que com duas fontes.

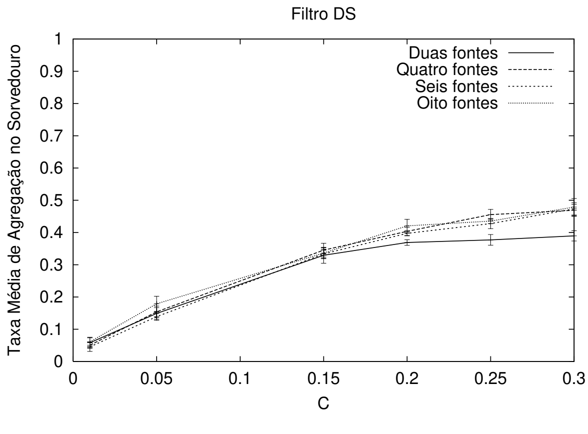


Figura 5.5: Variação do grau de certeza das proposições.

Filtro DS								
	2 fontes		4 fontes		6 fontes		8 fontes	
V	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95
0,10	14650,70	1147,12	35822,10	2317,78	57165,90	3851,76	83825,20	5781,93
0,20	14078,30	1129,89	37293,20	1963,39	62999,30	6121,30	88300,30	3388,15
0,30	14224,40	762,96	37493,50	2828,02	66998,40	3099,99	90555,00	4344,13
0,40	14018,80	981,22	37765,90	1558,49	65113,60	4376,83	88988,80	6102,41
0,50	13615,00	695,33	37835,10	2285,44	62830,60	6110,83	88397,90	4946,74
0,60	12451,00	1053,32	34463,50	2503,85	60867,10	2625,56	89630,30	4873,18
0,70	12839,80	619,15	34631,40	2043,39	62011,50	2830,56	83108,30	5791,29
0,80	12136,50	1177,97	33987,40	2196,21	60531,70	5176,55	87393,30	5215,59
0,90	11620,10	755,98	33812,50	3408,56	60339,30	3314,00	88530,50	5246,44
1,00	11534,80	595,15	33812,50	3408,56	59925,10	2840,96	83679,00	3094,18
1,10	12740,60	762,14	36653,90	1485,81	60212,50	3289,04	80484,00	3156,13
1,20	12621,90	622,52	34087,40	2313,60	59728,90	3789,23	85964,50	5198,01
	2 fontes		4 fontes		6 fontes		8 fontes	
D	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95
0,10	15547,70	916,60	45396,00	2116,33	72187,00	5560,54	103565,60	7160,54
0,20	15208,50	1133,50	44704,00	1567,81	73244,80	3342,30	100602,20	6737,97
0,30	15241,60	1082,68	42302,30	2779,82	71591,50	4024,67	100321,60	2885,25
0,40	14917,70	629,25	40888,50	3639,37	69739,40	3709,23	95638,30	5843,05
0,50	14331,80	848,82	39498,30	2132,29	67003,80	2671,65	96840,30	6155,99
0,60	13723,30	1091,86	39741,30	1985,07	65354,30	3435,36	88943,40	6435,14
0,70	13476,30	1047,14	41550,60	2080,80	65626,50	4978,62	88580,50	7766,08
0,80	14186,90	865,04	39580,00	1886,70	62888,60	3250,09	93943,10	5091,15
0,90	12630,10	843,44	34988,30	2212,14	61231,90	3224,06	90280,90	3770,26
1,00	12556,50	579,97	34736,70	1912,44	61715,40	4024,66	86864,30	3261,62
1,10	11319,50	900,12	33506,90	3543,04	58789,80	3032,92	84515,90	3264,54
1,20	10700,00	642,65	33597,60	2358,08	58212,90	2337,96	79655,30	4418,19

Tabela 5.2: Número de saltos no Filtro DS (com variação do parâmetro V e D).

Filtro DS								
	2 fontes		4 fontes		6 fontes		8 fontes	
C	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95	Saltos	IC-95
0,01	15467,60	897,90	45981,50	3058,20	77265,00	4392,81	111437,30	7091,84
0,05	13840,10	758,64	39944,20	3647,00	70575,20	2120,41	95822,60	6452,08
0,10	11500,70	1121,02	34983,50	2416,01	58969,50	4385,30	83129,90	2643,04
0,15	10489,00	895,16	31075,90	1725,53	51175,20	2582,73	76513,50	2753,38
0,20	9808,70	429,03	28198,30	1529,50	49289,90	2429,63	68932,50	3218,07
0,25	9976,20	633,62	24606,30	1174,46	45410,20	1985,19	63114,00	1563,79
0,30	9183,80	604,56	25066,00	1576,52	40978,90	2206,48	58686,90	3247,51

Tabela 5.3: Número de saltos no Filtro DS (com variação do parâmetro C).

O valor estabelecido para V, D e C deve ser ajustado em função dos requisitos da aplicação. Para os cenários seguintes foi adotado $T_v = I_t$ e $T_d = I_t$, ou seja, $V = 1,0$ e $D = 1,0$. No caso do parâmetro C, foi definido em 10% a variação máxima do grau de certeza das proposições dos sensores.

5.2.2 Cenário 2: Escalabilidade

Esse cenário avalia o comportamento da rede com o número de nós variando de 100 a 400. Os nós estão distribuídos em uma topologia em grade N x N (com N variando de 10 a 20 nós) e seis fontes são posicionadas duas a duas nas extremidades do campo de sensoriamento, como ilustrado na figura 5.6. Esta distribuição das fontes explora a situação em que o fenômeno de interesse é observado distante do sorvedouro, permitindo a formação de pontos de agregação.

A figura 5.7 compara o resultado da aplicação dos Filtros Bayes e DS fazendo a agregação, e da Difusão Direcionada pura. Os gráficos das figuras 5.7(a) e 5.7(b) indicam que o aumento da quantidade de nós na rede não teve influência significativa no desempenho da Difusão Direcionada com e sem a agregação, embora seja evidente a redução do número de mensagens no sorvedouro com a utilização de ambos os filtros, para qualquer das grades utilizadas.

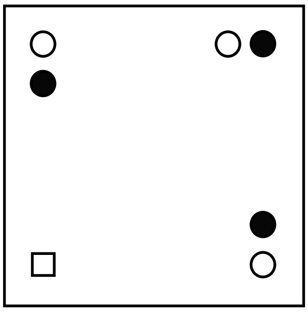
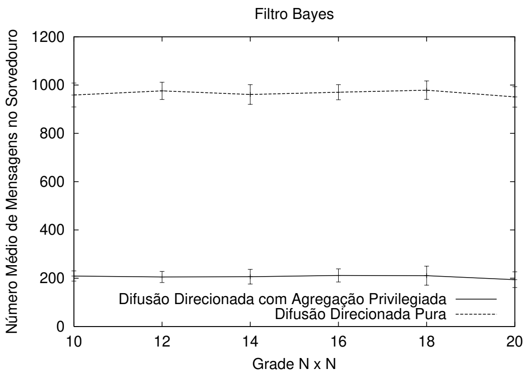
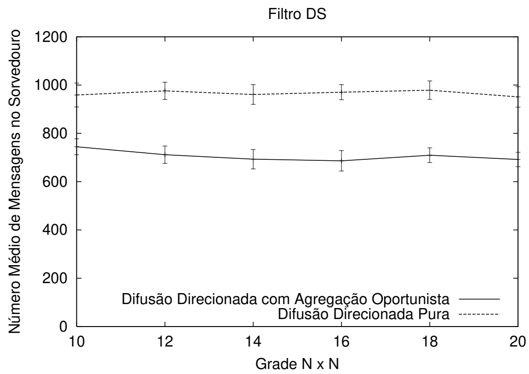


Figura 5.6: Posição das fontes.



(a) Agregação com o Filtro Bayes.



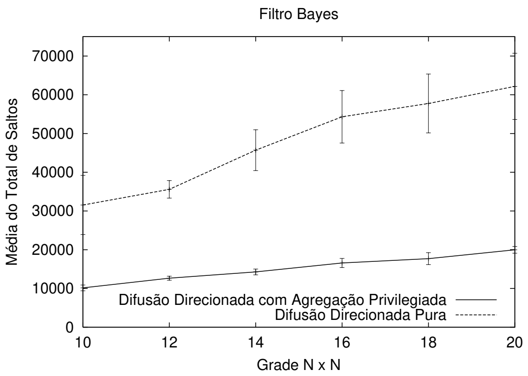
(b) Agregação com o Filtro DS.

Figura 5.7: Mensagens no Sorvedouro.

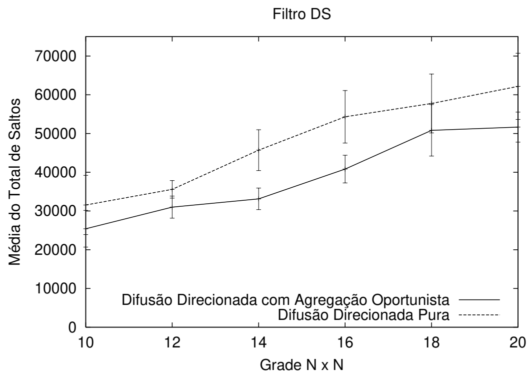
O acréscimo do número de transmissões mostrado nos gráficos da figura 5.8, pode ser explicado pelo aumento do comprimento das rotas entre as fontes e o sorvedouro. No gráfico da figura 5.8(a) pode ser observado que o desempenho do Filtro Bayes sofre pouca variação com o aumento da grade, mantendo o número de transmissões, aproximadamente, 65 a 70% menor do que a Difusão Direcionada pura. O mesmo acontece com o Filtro DS, cujo o número de transmissões foi, aproximadamente, 13 a 25% menor do que o roteamento sem agregação.

5.2.3 Cenário 3: Número de Fontes

Nesse cenário o número de fontes varia de dois a doze, sendo posicionadas em uma área quadrada com dez nós de lado. A distribuição das fontes explora uma



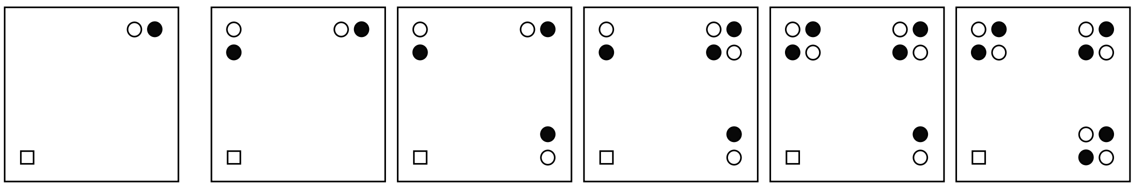
(a) Agregação com o Filtro Bayes.



(b) Agregação com o Filtro DS.

Figura 5.8: Números de saltos.

situação análoga a do cenário 2, como mostra a figura 5.9. Como era esperado, o aumento do número de fontes eleva a quantidade de mensagens no sorvedouro e também o número de saltos na rede.

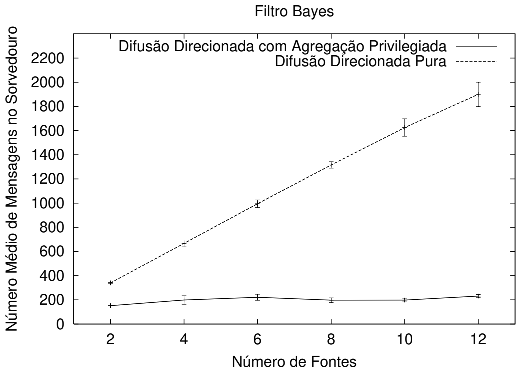


(a) 2 fontes. (b) 4 fontes. (c) 6 fontes. (d) 8 fontes. (e) 10 fontes. (f) 12 fontes.

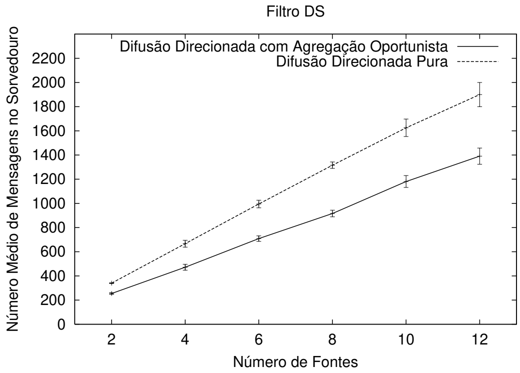
Figura 5.9: Topologia regular com variação do número de fontes.

A figura 5.10 mostra os resultados obtidos com as simulações do Filtro Bayes e do Filtro DS com relação ao número de mensagens recebidas no sorvedouro. Na simulação com o Filtro Bayes (figura 5.10(a)), observa-se uma redução drástica do número de mensagens que são entregues no sorvedouro, na medida em que cresce o número de fontes. Para duas fontes a redução é da ordem de 40%, enquanto com doze fontes ultrapassa 80%, quando comparada a Difusão Direcionada sem agregação. Com o Filtro DS a agregação mantém essa redução em aproximadamente 30% com a variação das fontes.

O número total de transmissões também foi avaliada nos dois filtros e ilustrada



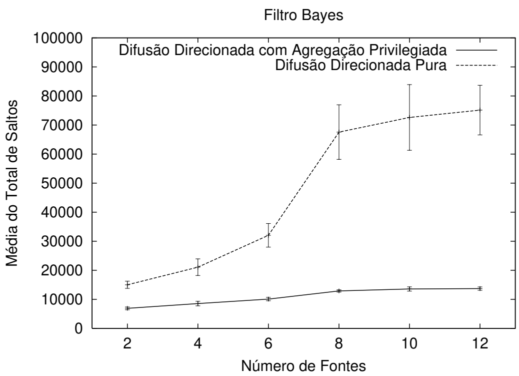
(a) Agregação com o Filtro Bayes.



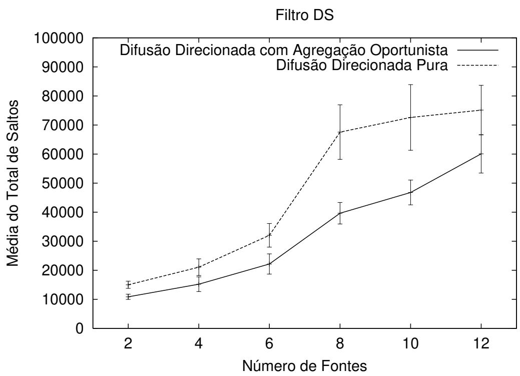
(b) Agregação com o Filtro DS.

Figura 5.10: Mensagens no sorvedouro.

na figura 5.11. No gráfico da figura 5.11(a) pode ser observado que o desempenho do Filtro Bayes melhora com o aumento do número de fontes. Com duas fontes a redução é de cerca de 55%, enquanto que, com doze fontes, chega a ordem de 80% em relação a Difusão Direcionada pura. Com o Filtro DS a redução de transmissões também é significativa, no entanto a maior redução foi de aproximadamente 42%, obtida com oito fontes.



(a) Agregação com o Filtro Bayes.

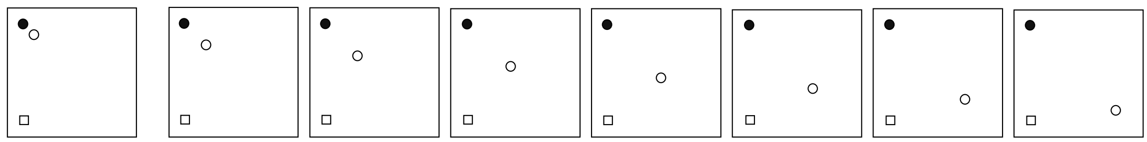


(b) Agregação com o Filtro DS.

Figura 5.11: Números de saltos.

5.2.4 Cenário 4: Posição Relativa das Fontes

O cenário 4 avalia a agregação em função da posição relativa entre as fontes. Para tal, foi mantida a topologia em grade com 100 nós, na qual foram posicionadas duas fontes eqüidistantes do sorvedouro em número de saltos e com uma separação mínima entre elas que varia entre dois e dezesseis saltos. A topologia do cenário é ilustrada na figura 5.12.



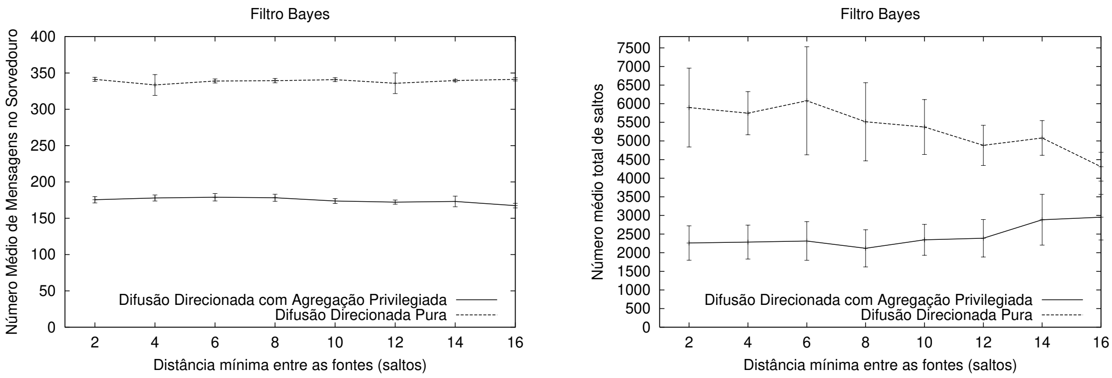
(a) 2. (b) 4. (c) 6. (d) 8. (e) 10. (f) 12. (g) 14. (h) 16.

Figura 5.12: Topologia com fontes eqüidistantes do sorvedouro e com separação mínima de s saltos entre si (s variando entre 2 e 16).

A posição relativa das fontes é importante para a eficiência da agregação, pois influencia a formação dos pontos onde ela é aplicada. Os gráficos da figura 5.13 apresentam o resultado da aplicação do Filtro Bayes comparada com a Difusão Direcionada pura. O gráfico da figura 5.13(a) indica que a redução de mensagens no sorvedouro teve uma pequena variação nas oito situações avaliadas (entre 47 e 52%). Entretanto, a redução no número de transmissões variou em cerca de 30% dependendo da posição entre as fontes, como mostra o gráfico da figura 5.13(b).

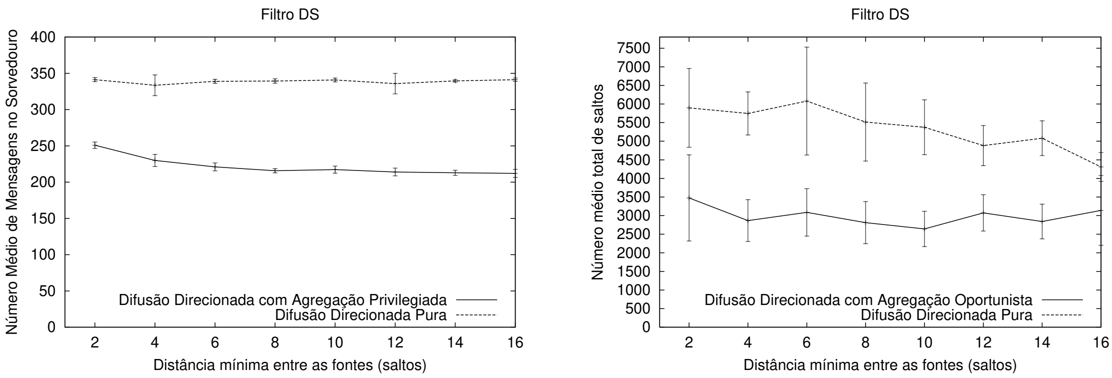
A aplicação do Filtro DS, no mesmo cenário, é apresentado nos gráficos da figura 5.14. Nesse caso, também foi pequena a variação da redução do número de mensagens no sorvedouro (entre 31 e 35%), como indica o gráfico da figura 5.14(a), enquanto a redução do número de transmissões variou entre 21 e 40 %, como ilustrado no gráfico da figura 5.14(b).

Dos resultados obtidos, pode-se inferir que as fontes posicionadas próximas entre si, guardada a mesma distância do sorvedouro, favorecem a aplicação do Filtro Bayes, resultando em uma redução mais expressiva no número de transmissões para uma quantidade semelhante de mensagens entregues no sorvedouro. A mesma ten-



(a) Número de mensagens no sorvedouro. (b) Número de saltos.

Figura 5.13: Variação da distância mínima das fontes em relação ao sorvedouro.



(a) Número de mensagens no sorvedouro. (b) Número de saltos.

Figura 5.14: Variação da distância mínima das fontes em relação ao sorvedouro.

dência não foi observada com o Filtro DS, embora as reduções de mensagens no sorvedouro e de transmissões também tenham sido significativas.

5.2.5 Cenário 5: Topologia Aleatória

Neste cenário, os filtros de agregação são avaliados em uma topologia com os nós distribuídos aleatoriamente. O campo de sensoriamento ocupou um quadrado de mil metros de lado, onde foram distribuídos os sensores. O evento de interesse ocorre na extremidade oposta ao sorvedouro, como ilustrado na figura 5.15. As simulações foram realizadas nas seguintes situações: para um número fixo de nós sensores, variando o número de fontes que detecta o evento entre 5 e 10% do total de nós sensores; e para um número fixo de fontes, variando o total de nós no campo de sensoriamento entre 100 e 200.

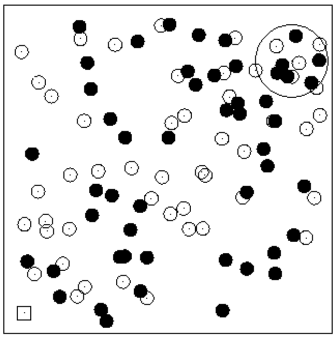
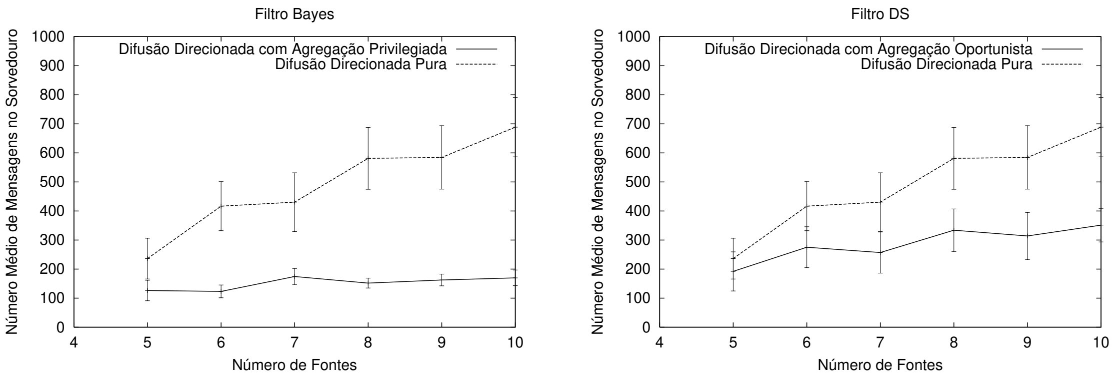


Figura 5.15: Topologia aleatória.

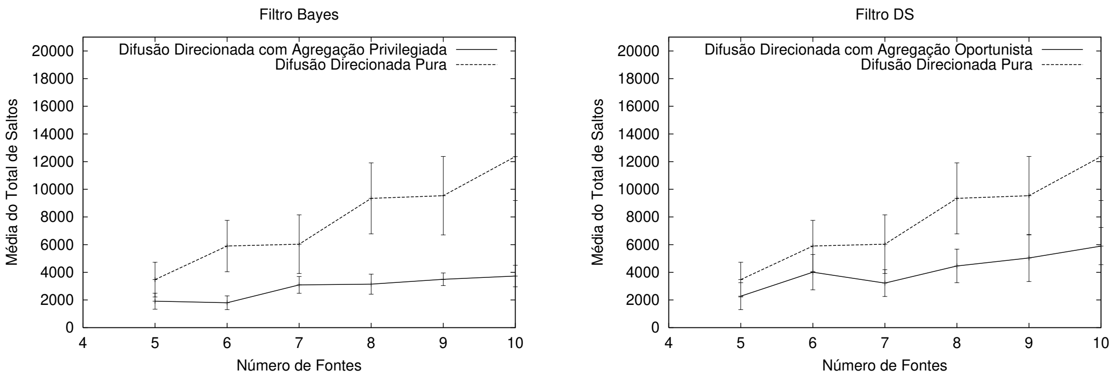
Os gráficos da figura 5.16 mostram que ambos os filtros conseguem reduzir o número de mensagens no sorvedouro. O Filtro Bayes tem um desempenho mais favorável a medida que cresce o número de fontes (figura 5.16(a)), pois mantém aproximadamente o mesmo número de entrega de mensagens no sorvedouro, mesmo com o aumento de fontes, o que não ocorre com a Difusão Direcionada pura. No Filtro Bayes com cinco fontes, o número de mensagens é aproximadamente 47% menor, enquanto que, com dez fontes, essa redução chega a 76%. Na agregação com o Filtro DS, o aumento do número de fontes também melhora seu desempenho (figura 5.16(b)). Nesse caso, o número de mensagens que chega no sorvedouro com cinco fontes corresponde a aproximadamente 81% do total entregue com o roteamento sem agregação, enquanto que com dez fontes esse valor fica próximo de 51%.

A aplicação dos filtros de agregação com a distribuição aleatória dos nós, também obteve um desempenho favorável com relação ao número de transmissões a medida que o número de fontes aumenta, como ilustrado na figura 5.17. O Filtro Bayes (figura 5.17(a)) obteve a maior redução de transmissões com dez fontes (aproximadamente 70%) e a menor com cinco fontes (aproximadamente 45%). No Filtro DS (figura 5.17(b)) a redução foi menos expressiva, variando aproximadamente entre 32% (com seis fontes) e 53% (com dez fontes).



(a) Agregação com o Filtro Bayes. (b) Agregação com o Filtro DS.

Figura 5.16: Mensagens no sorvedouro com a variação do número de fontes.

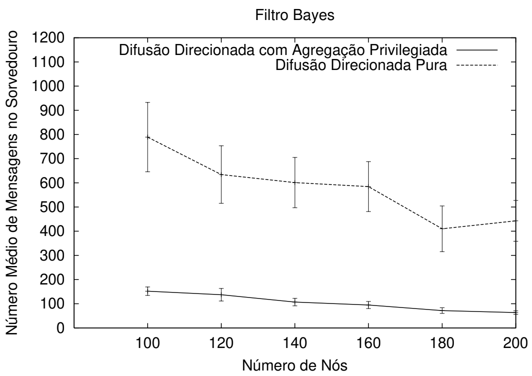


(a) Agregação com o Filtro Bayes. (b) Agregação com o Filtro DS.

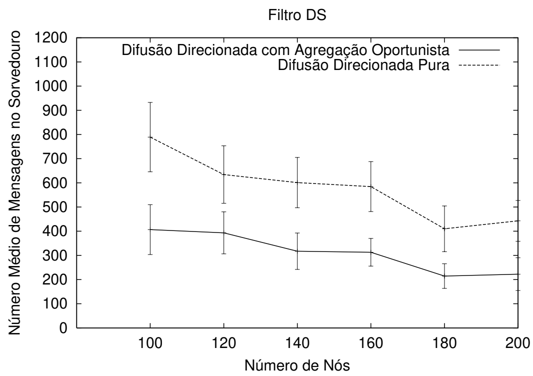
Figura 5.17: Números de saltos com a variação do número de fontes.

Os gráficos das figuras 5.18 e 5.19 mostram o comportamento dos filtros de agregação com o aumento da densidade dos nós em uma topologia aleatória. O Filtro Bayes (figura 5.18(a)) apresentou seu melhor resultado com duzentos nós, o que corresponde a uma redução de 86% mensagens entregues no sorvedouro em relação a Difusão Direcionada pura. O mesmo resultado foi obtido com o Filtro DS (figura 5.18(b)), sendo que a redução foi aproximadamente de 50%.

Com relação ao número de transmissões, o Filtro Bayes teve um pequeno favorecimento com o aumento do número de nós (figura 5.19(a)), mantendo o total de transmissões aproximadamente 80% menor do que o roteamento sem agregação, enquanto que o Filtro DS não apresentou uma tendência definida (figura 5.19(b)), obtendo praticamente o mesmo percentual de redução de transmissões com cem e duzentos nós.

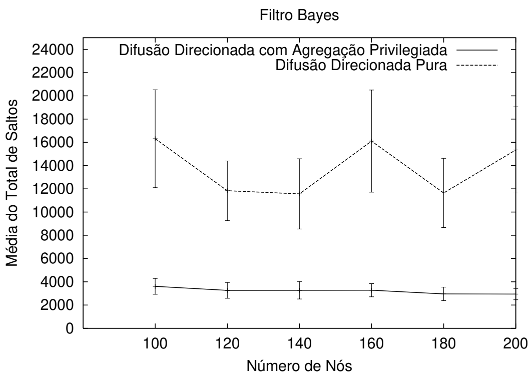


(a) Agregação com o Filtro Bayes.

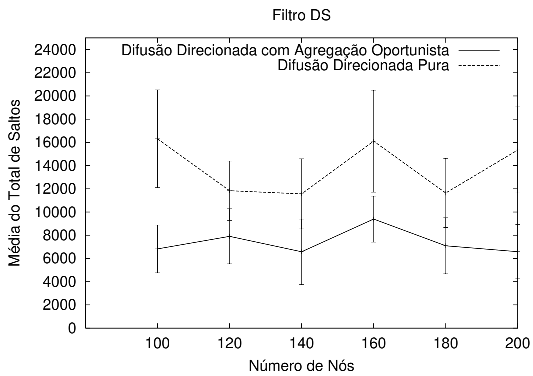


(b) Agregação com o Filtro DS.

Figura 5.18: Mensagens no sorvedouro com a variação do número de nós.



(a) Agregação com o Filtro Bayes.



(b) Agregação com o Filtro DS.

Figura 5.19: Números de saltos com a variação do número de nós.

5.2.6 Cenário 6: Latência da Rede

Neste cenário é avaliado o impacto na latência da rede comparando-se o atraso total dos pacotes entre a fonte e o sorvedouro, desprezando-se o tempo de processamento gasto nos filtros envolvidos. Foi utilizada a topologia em grade com cem nós e com duas fontes posicionadas na extremidade oposta ao sorvedouro, como ilustrado na figura 5.20. Os resultados foram obtidos com a Difusão Direcionada sem realizar qualquer tipo de agregação, com a aplicação do Filtro DS que implementa o mecanismo de agregação proposto e com o Filtro TAF (Filtro de Agregação Tradicional), que foi desenvolvido para este cenário, com base no esquema de armazenar e reencaminhar mensagens. As três situações simuladas foram realizadas com o estabelecimento de rotas que priorizam a agregação, ou seja, em conjunto com o Filtro PAR.

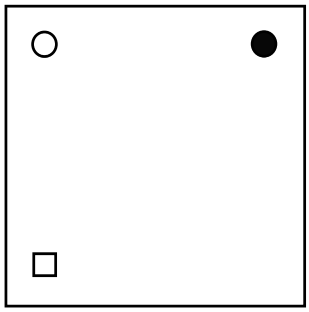


Figura 5.20: Posição do sorvedouro e das fontes.

O Filtro TAF realiza agregação em cada nó selecionando a mensagem com menor grau de incerteza recebida durante um tempo de espera (T_e) pré-definido. Como não é possível determinar antecipadamente o ponto de agregação, cada nó emprega um *cache* associado a um temporizador que retém as mensagens de dados recebidas por um período T_e. Após esse tempo, é aplicado o critério de seleção para a escolha da mensagem que será encaminhada para o próximo nó, descartando as demais. Como no cenário considerado só existem duas fontes, o filtro foi configurado para realizar no máximo uma agregação. Portanto, uma mensagem que sofre uma agregação não terá que esperar um novo período T_e no caminho fonte-sorvedouro.

Os resultados obtidos neste cenário são mostrados na tabela 5.4. Cada linha numerada de 1 a 10 mostra o resultado de uma rodada de simulação. As duas últimas linhas apresentam, respectivamente, a média e o intervalo de confiança de 95% para as 10 rodadas de simulação.

Nessa tabela, a coluna "Difusão Direcionada"mostra o atraso médio de cada pacote entregue no sorvedouro. A coluna "Filtro DS"contabiliza o atraso médio dos

	Difusão Direcionada	Filtro DS	Filtro TAF				
			T_e (ms)				
			10	20	40	80	160
Rodada	Retardo (ms)						
1	31,49	33,74	137,87	211,62	430,39	730,04	1194,88
2	30,05	39,47	133,69	224,06	433,24	902,51	1136,32
3	39,16	34,41	137,90	264,98	330,18	701,14	1292,12
4	40,41	38,30	115,23	199,73	507,16	601,95	1419,73
5	36,92	37,86	108,38	289,78	336,79	811,77	1625,34
6	34,73	33,97	118,79	241,53	476,94	832,31	1402,00
7	31,10	32,80	137,09	191,51	389,72	777,04	1257,73
8	40,41	39,97	133,59	226,30	434,38	889,82	1812,54
9	37,48	30,26	131,59	272,53	512,32	704,07	1827,13
10	37,89	37,93	160,56	231,92	387,54	636,08	1149,29
média	35,96	35,87	131,47	235,40	423,86	758,67	1411,71
IC 95	2,79	2,33	10,47	22,91	45,66	72,78	186,04

Tabela 5.4: Impacto na latência da rede.

pacotes com o emprego do mecanismo de agregação proposto realizando a Inferência de Dempster-Shafer. Como é possível verificar na tabela, não existe diferença significativa para os dois casos, uma vez que a agregação realizada pelo Filtro DS ocorre sem tempo de espera.

A coluna "Filtro TAF"mostra a média das amostras de retardo para diversos valores de T_e. O valor inicial de T_e corresponde ao menor período necessário para que ocorra a agregação no cenário avaliado. Os demais valores foram selecionados dobrando o valor T_e até 160 milissegundos. O resultado obtido é coerente com o esperado, conforme discutido na seção 3.2.1. Na tabela 5.4, verifica-se que mesmo para $T_e = 10$ milissegundos, o atraso médio dos pacotes é superior a três vezes o atraso provocado com o Filtro DS.

		Influência dos parâmetros do mecanismo de agregação			
		Filtro Bayes (agregação privilegiada)		Filtro DS (agregação oportunista)	
		Mensagens no Sorvedouro	Número de Saltos	Mensagens no Sorvedouro	Número de Saltos
cenário 1	V	$\downarrow (0,9)$	$\nearrow$	$\downarrow (0,9)$	$\nearrow$
(2 fontes)	D	$\downarrow (1,0)$	$\nearrow$	$\downarrow (1,0)$	$\nearrow$
parâmetros	C	-	-	$\nearrow$	$\nearrow$

$\downarrow (n)$ - indica o valor do parâmetro ($n \leq 1,0$) que obteve a maior redução

$\nearrow$ - indica que o aumento do parâmetro favorece a métrica de desempenho

Tabela 5.5: Resumo dos parâmetros do mecanismo para duas fontes.

	Redução aproximada em relação à Difusão Direcionada pura							
	Filtro Bayes (agregação privilegiada)				Filtro DS (agregação oportunista)			
	Mensagens no Sorvedouro		Número de Saltos		Mensagens no Sorvedouro		Número de Saltos	
	Menor	Maior	Menor	Maior	Menor	Maior	Menor	Maior
cenário 2 (número de nós)	78% (100)	80% (400)	65% (144)	70% (324)	22% (100)	29% (256)	13% (144)	25% (256)
cenário 3 (número de fontes)	40% (2)	80% (12)	55% (2)	80% (12)	25% (2)	30% (8)	21% (12)	42% (8)
cenário 4 (distância das fontes)	47% (4)	52% (16)	32% (16)	62% (2)	31% (2)	35% (16)	21% (4)	40% (6)
cenário 5 (número de fontes)	47% (5)	76% (10)	45% (5)	70% (10)	19% (5)	49% (10)	32% (6)	53% (10)
(número de nós)	78% (120)	86% (200)	72% (120)	81% (200)	38% (120)	50% (200)	33% (120)	58% (100)

Tabela 5.6: Resumo dos resultados obtidos nos cenários de 2 a 5.

5.2.7 Resumo dos Resultados

O resumo dos resultados obtidos nas simulações é apresentado nesta seção. A tabela 5.5 mostra os valores dos parâmetros que obtiveram a maior redução do número de mensagens no sorvedouro e sua influência no número de saltos na rede. A tabela 5.6 traz os maiores e menores valores percentuais de cada métrica de desempenho, que foram obtidos nos cenários dois a cinco. E finalmente, a tabela 5.7 resume o impacto na latência da rede simulado no cenário seis. Com os resultados obtidos, é possível concluir que o mecanismo de agregação obteve uma economia de energia significativa para todos os cenários avaliados, com um baixo impacto na rede, particularmente, quanto ao atraso na entrega das mensagens.

	Difusão Direcionada	Filtro DS	Filtro TAF				
	Atraso médio (ms)						
Cenário 6	35,96	35,87	131,47	235,40	423,86	758,67	1411,71
(T_e ms)	-	(0)	(10)	(20)	(40)	(80)	(160)

Tabela 5.7: Resumo do impacto na latência da rede.

5.3 Conclusão Parcial

Este capítulo abordou a avaliação de desempenho do mecanismo de agregação empregando as técnicas paramétricas de Inferência Bayesiana e de Dempster-Shafer, utilizando o simulador *ns-2* [9]. Para tal, foram adicionados os Filtros Bayes e DS ao módulo da Difusão Direcionada, de forma a explorar a formação de pontos de agregação privilegiados (com o Filtro PAR) e oportunistas (Difusão Direcionada pura). Ambos os filtros de agregação foram implementados de forma a minimizar os custos da agregação em RSSFs.

Os resultados dos cenários avaliados mostraram, segundo as simulações realizadas, que o mecanismo de agregação proposto reduz drasticamente o número de mensagens recebidas no sorvedouro, assim como, o número total de transmissões necessárias para a execução das tarefas de sensoriamento na rede, sem a introdução de atraso significativo na entrega das mensagens.

O próximo capítulo apresenta as conclusões da dissertação e as sugestões de trabalhos futuros.

Capítulo 6

Conclusões e Trabalhos Futuros

O objetivo deste trabalho de dissertação foi o desenvolvimento de um mecanismo de agregação de dados empregando técnicas paramétricas em RSSFs. Tal mecanismo tem por finalidade proporcionar economia de energia por meio da redução da comunicação entre os nós sensores e, ao mesmo tempo, refinar as informações de interesse, minimizando os impactos na latência e robustez da rede.

As principais contribuições deste trabalho foram:

- a proposta de um mecanismo de agregação que difere do procedimento tradicional de armazenar e reencaminhar mensagens;
- a aplicação e adaptação da Inferência Bayesiana e de Dempster-Shafer para agregação de dados em RSSFs, explorando o modelo JDL DFS de fusão de dados e as características do protocolo de roteamento centrado em dados; e
- os resultados das simulações que mostraram, para os cenários avaliados, uma economia de energia significativa com um baixo custo na rede.

A próxima seção apresenta as conclusões do trabalho e a seção 6.2 traz sugestões para trabalhos futuros motivados a partir da pesquisa desenvolvida neste trabalho.

6.1 Conclusões

As RSSFs possuem um elevado número de nós que combinam as capacidades de sensoriamento e de processamento, utilizam o meio sem fio para comunicação e dependem da energia economizada em cada aspecto de sua arquitetura para determinar seu tempo de vida. Sendo a comunicação entre os nós a maior responsável pelo consumo de energia, o aumento do processamento dos dados em cada nó consiste em uma alternativa para economizar energia. Esse processamento realizado de forma distribuída é explorado pela agregação de dados durante o roteamento das mensagens na rede.

Dentre os protocolos de roteamento para RSSFs propostos na literatura, a Difusão Direcionada é um protocolo robusto e de relativa simplicidade, requisito essencial considerando-se os escassos recursos dos sensores. O mecanismo de agregação proposto foi implementado tendo como base esse protocolo e buscando preservar suas características como, por exemplo, as interações localizadas e a independência da topologia, na sua interação com os demais componentes da arquitetura. Outra característica importante desse protocolo é ser centrado em dados, isto é, ele permite um roteamento baseado em atributos, o que favorece a agregação empregando técnicas paramétricas.

As técnicas paramétricas selecionadas para a agregação de dados foram a Inferência Bayesiana e de Dempster-Shafer, que tornam possível realizar a agregação de diferentes tipos de dados oriundos de múltiplos sensores em aplicações de detecção, identificação e classificação de eventos. Estas técnicas permitem o mapeamento direto entre o fenômeno que se quer observar e as evidências detectadas pelos sensores, que caracterizam tal fenômeno. Embora tais técnicas tenham sido alvos de diversas pesquisas e propostas para fusão de dados de sensores tradicionais como, por exemplo, radares, imageadores orbitais ou embarcados em aeronaves, a literatura em RSSFs tem se limitado a citar o seu potencial, sem apresentar uma proposta que minimize os impactos da agregação na rede.

O mecanismo proposto buscou reduzir os impactos da agregação em RSSFs introduzidos pelos esquemas tradicionais, em particular, o retardo devido ao atraso dos dados oriundos de fontes mais afastadas do ponto de agregação. Tal mecanismo foi implementado nos Filtros Bayes e DS. Ambos os filtros atuam de forma descentralizada em uma RSSF para realizar a agregação privilegiada (com o emprego do Filtro PAR) e oportunista (com a Difusão Direcionada pura), eliminando dados redundantes e inferindo antecipadamente (antes do sorvedouro) sobre o fenômeno de interesse. A agregação é realizada com base nas observações do nó sensor e nas informações obtidas dos nós vizinhos, no momento em que recebe uma mensagem, sem adicionar ciclos de espera.

A definição dos parâmetros do mecanismo de agregação é uma tarefa complexa. Os valores ideais de tais parâmetros dependem dos requisitos da aplicação. No entanto, as simulações mostraram que os tempos de validade e de descarte podem ser configurados com o mesmo valor do menor intervalo de transmissão das fontes, de modo a obter uma economia de energia significativa, sem incorrer em profundas modificações no comportamento da rede. A variação máxima admitida para o grau de certeza, no caso do Filtro DS, explora o compromisso entre economia de energia e a precisão das proposições.

Na avaliação dos diversos cenários realizados sobre o consumo de energia na rede, foi verificado que o mecanismo de agregação proposto conseguiu atingir seus objetivos. Os resultados mostraram uma redução significativa das mensagens que chegam ao sorvedouro, assim como, o número total de transmissões realizadas para a execução das tarefas de sensoriamento.

Como observação final, embora o objetivo deste trabalho não inclua uma comparação entre as técnicas de agregação, pode-se destacar como desvantagens da Inferência Bayesiana, a dificuldade para definir as probabilidades a *priori* das hipóteses, o incremento da complexidade com o aumento das hipóteses e eventos dependentes, a exigência de hipóteses mutuamente exclusivas e a incapacidade de tratar a incerteza. Enquanto a Inferência de Dempster-Shafer resolve parte desses problemas, exige a definição de um algoritmo em cada sensor para a atribuição do grau de certeza da proposição.

6.2 Trabalhos Futuros

Como trabalhos futuros, sugere-se a pesquisa de outras técnicas paramétricas do modelo JDL DFS para a agregação de dados em RSSFs como, por exemplo, a Inferência Clássica e o Processamento Generalizado de Evidências (GEP), que não foram avaliadas neste trabalho. Estas técnicas podem ser empregadas em aplicações similares as que foram estudadas, envolvendo a detecção, classificação e a identificação de eventos. O estudo pode incluir, ainda, uma análise comparativa entre as técnicas aplicadas.

Quanto aos resultados das simulações, buscou-se estudar, vários cenários para as métricas de desempenho selecionadas, segundo os objetivos do trabalho. No entanto, existem outros cenários onde a agregação de dados em RSSFs empregando técnicas paramétricas necessitam ser avaliadas. Por exemplo, os cenários com mudanças de topologia, envolvendo mobilidade ou falhas nos nós motivadas pelo esgotamento de energia, interferências ou destruição. Estes cenários introduzem novas métricas de desempenho, como o consumo e a distribuição de energia na rede, merecendo uma pesquisa futura.

O uso de um exemplo genérico neste trabalho foi com o objetivo de não ficar preso a uma aplicação específica, ampliando o potencial de emprego da proposta. Por isso, a implementação de aplicações específicas, envolvendo modelos com várias hipóteses e evidências, é sugerida como forma de indicar quais as aplicações proporcionam os melhores resultados e mensurar o nível de complexidade adicionado.

Outra sugestão de trabalho futuro é a integração em um único módulo das funcionalidades existentes no Filtro PAR e nos Filtros Bayes ou DS, tirando proveito da arquitetura modular da Difusão Direcionada implementada no simulador *ns-2*. Dessa forma, o mesmo filtro seria responsável pela seleção de rotas que forcem a passagem dos fluxos de dados por pontos comuns e a agregação dos dados.

Referências Bibliográficas

[1] WATLZ, E., AND LLINAS, J. *Multisensor Data Fusion.* Artech House, 1990.

[2] GONÇALVES, P. A. S., DE REZENDE, J. F., AND DUARTE, O. C. M. B. Um serviço ativo de distribuição de vídeo multiponto. *18º Simpósio Brasileiro de Redes de Computadores (SBRC 2000), Belo Horizonte, MG, Brasil* (2000).

[3] KRISHNAMACHARI, B., ESTRIN, D., AND WICKER, S. Modelling data-centric routing in wireless sensor networks. *The 21st Annual Joint Conference of the IEEE Computer Communications Societies (INFOCOM 2002), New York, NY, USA* (junho de 2002).

[4] INTANAGONWIWAT, C., ESTRIN, D., GOVINDAN, R., AND HEIDEMANN, J. Impact of network density on data aggregation in wireless sensor networks. *Proceedings of the 22nd International Conference on Distributed Computing Systems* (julho de 2002).

[5] KRISHNAMACHARI, B., ESTRIN, D., AND WICKER, S. The impact of data aggregation in wireless sensor networks. *International Workshop on Distributed Event-Based Systems, (DEBS '02)* (julho de 2002).

[6] YUAN, W., KRISHNAMURTHY, S. V., AND TRIPATHI, S. K. Synchronization of multiple levels of data fusion in wireless sensor networks. *IEEE Global Communications Conference (GLOBECOM 2003), San Francisco, CA, USA* (dezembro de 2003).

[7] INTANAGONWIWAT, C., GOVINDAN, R., AND ESTRIN, D. Directed diffusion: A scalable and robust communication paradigm for sensor networks.

Proceedings of 6th Annual International Conference on Mobile Computing and Networking (MOBICOM-00) (agosto de 2000), 56–67.

[8] Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., and Silva, F. Directed diffusion for wireless sensor networking. *IEEE/ACM Transactions on Networking (TON) 11*, 1 (fevereiro de 2002), 2–16.

[9] The Network Simulator ns-2. Página da World Wide Web. url: `http://www.isi.edu/nsnam/ns/`, março de 2002.

[10] Pinto, A. J. G., Stochero, J. M., and de Rezende, J. F. Descoberta de rotas e mecanismo de agregação em redes de sensores. *22º Simpósio Brasileiro de Redes de Computadores (SBRC 2004), Gramado, RS, Brasil* (2004).

[11] Pinto, A. J. G., Stochero, J. M., and de Rezende, J. F. Proposta e avaliação de um mecanismo de agregação de dados em redes de sensores. *A ser publicado em: III WPerformance / Workshop em Desempenho de Sistemas Computacionais e de Comunicação / XXIV Congresso da Sociedade Brasileira de Computação (SBC 2004), Salvador, BA, Brasil* (2004).

[12] Internet Engineering Task Force (IETF). Mobile ad-hoc networks (MANET) working group. Página da World Wide Web. url: `http://www.ietf.org/html.charters/manet-charter.html/`, janeiro de 2004.

[13] Tanenbaum, A. S. *Computers Networks - Third Edition*. Pretence-Hall, Inc., 1979.

[14] Telecommunications standardization sector. Página da World Wide Web. url: `http://www.itu.int/ITU-T/`, janeiro de 2004.

[15] Institute of Electrical and Eletronics Engineers (IEEE). Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. IEEE Santandard 802.11, 1999.

[16] Institute of Electrical and Eletronics Engineers (IEEE). 802.11 Wireless Local Area Networks. Página da World Wide Web. url: http://www.ieee802.org/11/, dezembro de 2003.

[17] Institute of Electrical and Eletronics Engineers (IEEE). Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-speed Physical Extension in the 5 GHz Band. IEEE Santandard 802.11a, 1999.

[18] Institute of Electrical and Eletronics Engineers (IEEE). Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: High-speed Physical Extension in the 2.4 GHz Band. IEEE Santandard 802.11b, 1999.

[19] Institute of Electrical and Eletronics Engineers (IEEE). IEEE 802.15 Working Group for Wireless Personal Area Networks. Página da World Wide Web. url: http://www.ieee802.org/15/, dezembro de 2003.

[20] Institute of Electrical and Eletronics Engineers (IEEE). IEEE 802.16 Working Group on Broadband Wireless Access Standards. Página da World Wide Web. url: http://www.ieee802.org/16/, dezembro de 2003.

[21] Institute of Electrical and Eletronics Engineers (IEEE). IEEE 802.20 Mobile Broadband Wireless Access. Página da World Wide Web. url: http://grouper.ieee.org/groups/802/20/, maio de 2004.

[22] Sohrabi, K., Gao, J., Ailawadhi, V., and Pottie, G. J. Protocols for self-organization of a wireless sensor network. *IEEE Personal Communications 7*, 5 (outubro de 2000), 16–27.

[23] Pottie, G. J., and Kaiser, W. J. Wireless integrated network sensors (WINS). *Communications of the ACM 43*, 5 (maio de 2000), 51–58.

[24] Mini, A. F., Nath, B., and Loureiro, A. A. F. Prediction-based approaches to construct the energy map for wireless sensor networks. *21º Simpósio Brasileiro de Redes de Computadores (SBRC 2003), Natal, RN, Brasil* (maio de 2003).

[25] Wong, J. L., and Potkonjak, M. Search in sensor networks: Challenges, techniques, and applications. *International Conference on Acoustics Speech and Signal Processing (ICASSP 2002), Orlando, FL, USA* (maio de 2002).

[26] Estrin, D., Girod, L., Pottie, G., and Srivastava, M. Instrumenting the world with wireless sensor networks. *In International Conference on Acoustics, Speech, and Signal Processing (ICASSP-01)* (maio de 2001).

[27] Feng, J., Koushanfar, F., and Potkonjak, M. System-architectures for sensor networks issues, alternatives, and directions. *International Conference on Computer Design (ICCD 2002), Freiburg, Germany* (setembro de 2002), 112–121.

[28] Eecs, F. K. Fault tolerance techniques for wireless ad hoc sensor networks. *The First IEEE International Conference on Sensors, Orlando, Florida, USA* (junho de 2002).

[29] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., and Cayirici, E. A survey on sensor networks. *Communications of the ACM* (agosto de 2002), 102–114.

[30] Bulusu, N., Estrin, D., Girod, L., and Heidemannn, J. Scalable coordination for wireless sensor networks: Self-configuring localization systems. *Proceedings of the International Symposium on Communication Theory and Applications (ISCTA)* (julho de 2001).

[31] Bluetooth Special Interest Group. Página da World Wide Web. url: `http://www.bluetooth.com/`, janeiro de 2004.

[32] Shah, R., and Rabaey, J. Energy aware routing for low energy ad hoc sensor networks. *Proceedings for the IEEE Wireless Communications and Networking Conference (WCNC 2002), Orlando, FL, USA* (março de 2002).

[33] Mini, R. A. F., Nath, B., and Loureiro, A. A. F. A probabilist approach to predict the energy consumption in wireless sensor networks. *IV Workshop de Comunicação sem Fio e Computação Móvel, São Paulo, SP* (outubro de 2002).

[34] Shih, E., Cho, S.-H., Ickes, N., Min, R., Sinha, A., Wang, A., and Chandrakasan, A. Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks. *The Seventh Annual International Conference on Mobile Computing and Networking (MobiCom 2001), Rome, Italy* (julho de 2001).

[35] Woo, A., and Culler, D. A transmission control scheme for media access in sensor networks. *In The Seventh Annual International Conference on Mobile Computing and Networking (MobiCom 2001), Rome, Italy* (julho de 2001), 221–235.

[36] Kahn, J., Katz, R., and Pister, K. Emerging challenges: Mobile networking for smart dust. *Journal of Communications and Networks* (setembro de 2000), 188–196.

[37] Zhou, L., and Haas, Z. J. Securing ad hoc networks. *IEEE Network Magazine 13*, 6 (novembro de 1999), 24–30.

[38] Perrig, A., Szewczyk, R., Wen, V., Culler, D. E., and Tygar, J. D. SPINS: security protocols for sensor netowrks. *The Seventh Annual International Conference on Mobile Computing and Networking (MobiCom 2001), Rome, Italy* (julho de 2001), 189–199.

[39] Yu, Y., Govindan, R., and Estrin, D. Geographical and energy aware routing: A recursive data dissemination protocol for wireless sensor networks. *UCLA Computer Science Department Technical Report UCLA/CSD-TR-01-0023* (maio de 2001).

[40] Elson, J., and Estrin, D. Time synchronization for wireless sensor networks. *Proceedings of the 15th International Parallel and Distributed Processing Symposium* (2001), 186.

[41] Raghunathan, V., Schurgers, C., Park, S., and Srivastava, M. Energy aware wireless microsensor networks. *IEEE Signal Processing Magazine 19*, 2 (março de 2002), 40–50.

[42] Hill, J., Szewczyk, R., Woo, A., Hollar, S., Culler, D., and Pister, K. System architecture directions for network sensors. *Proceedings of the 9th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS-00)* (novembro de 2000), 93–104.

[43] Shen, C.-C., Srisathapornphat, C., and Jaikaeo, C. Sensor information network architecture and application. *IEEE Personal Communication* (agosto de 2001).

[44] Heinzelman, W. R., Chandrakasan, A., and Balakrishnan, H. Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the 33rd IEEE Hawaii International Conference on System Sciences (HICSS), Maui, Hawaii, USA* (janeiro de 2000), 1–10.

[45] Asada, G., Dong, T., Lin, F., Pottie, G., Kaiser, W., and Marcy, H. Wireless integrated network sensors: Low power systems on a chip. *European Solid State Circuits Conference, The Hague, Netherlands.* (1998).

[46] Chien, C., Elgorriaga, I., and McConaghy, C. Low-power direct-sequence spread-spectrum modem architecture for distributed wireless sensor networks. *Proceedings of the International Symposium on Low Power Electronics and Design* (2001), 251–254.

[47] Mihai, B. Ultra wide band technologies. Ad Hoc Mobile Wireless Networks - Research Seninar on Telecomunications Software, Autumn 2002. HUT TML Course T-110.557 Publication ISBN 951-22-6309-2 ISSN 1456-7628 TML-C8.

[48] ZigBee Alliance. Página da World Wide Web. url: `http://www.zigbee.org/`, maio de 2004.

[49] Ye, W., Heidemann, J., and Estrin, D. An energy-efficient MAC protocol for wireless sensor networks. *The 21st Annual Joint Conference of the IEEE Computer Communications Societies (INFOCOM 2002), New York, NY, USA* (2002).

[50] Youssef, M., Younis, M., and Arisha, K. A constrained shortest-path energy-aware routing algorithm for wireless sensor networks. *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC 2002), Orlando, FL, USA* (março de 2001).

[51] Heinzelman, W., Kulik, J., and Balakrishnan, H. Adaptive protocols for information dissemination in wireless sensor networks. *Proceedings of 5th Annual International Conference on Mobile Computing and Networking (MobiCom '99)* (agosto de 1999), 174–185.

[52] Park, S., Savvides, A., and Srivastava, M. B. Simulating networks of wireless sensors. *Proceedings of 2001 Winter Simulation Conference (WSC 2001), Arlington, VA, USA* (dezembro de 2001), 1330–1338.

[53] Royer, E., and Toh, C. A review of current routing protocols for ad-hoc mobile wireless networks. *IEEE Personal Communications.* (abril de 1999).

[54] Rodoplu, V., and Meng, T. H. Minimum energy mobile wireless networks. *Proceedings of the IEEE International Conference on Communications (ICC '98) 3* (junho de 1998), 1633–1639.

[55] Li, L., and Halpern, J. Minimum energy mobile wireless networks revisited. *Proceedings of the IEEE International Conference on Communications (ICC 2001), St.Petersburg, Russia* (junho de 2001).

[56] Handy, M., Haase, M., and Timmermann, D. Low energy adaptive clustering hierarchy with deterministic cluster-head selection. *The Fourth IEEE Conference on Mobile and Wireless Communications Networks (MWCN 2002), Stockholm, Sweden* (setembro de 2002).

[57] Lindsey, S., and Raghavendra, C. S. Pegasis: Power-efficient gathering in sensor information systems. *Proceedings of the IEEE International Conference on Communications (ICC 2001), St.Petersburg, Russia* (2001).

[58] Sankarasubramaniam, Y., Özgür B. Akan, and Akyildiz, I. F. Esrt: event-to-sink reliable transport in wireless sensor networks. *Proceedings of the*

4th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc 2003), Annapolis, MD, USA (junho de 2003), 177–188.

[59] Stann, F., and Heidemann, J. Rmst: Reliable data transport in sensor networks. *Proceedings of the First International Workshop on Sensor Net Protocols and Applications, Anchorage, Alaska, USA* (abril de 2003), 102–112.

[60] Schwiebert, L., Gupta, S. K. S., Auner, P. S. G., Abrams, G., Iezzi, R., and McAllister, P. A biomedical smart sensor for the visually impaired. *IEEE Sensors 2002* (junho de 2002).

[61] Cerpa, A., Elson, J., Estrin, D., Girod, L., Hamilton, M., and Zhao, J. Habitat monitoring: Application driver for wireless communications technology. *Proceedings of the ACM SIGCOMM Workshop on Data Communications in Latin America and the Caribbean, San Jose, Costa Rica* (abril de 2001).

[62] Estrin, D., Govindan, R., Heidemann, J. S., and Kumar, S. Next century challenges: Scalable coordination in sensor networks. *Proceedings of 5th Annual International Conference on Mobile Computing and Networking (MobiCom '99), Seattle, Washington, USA* (agosto de 1999), 263–270.

[63] Nemeroff, J., Garcia, L., Hampel, D., and DiPietro, S. Application of sensor network communication. *Proceedings of the Communications for Network-Centric Operations (MILCOM 2001), Washington, DC, USA* (outubro de 2001).

[64] Park, S., Savvides, A., and Srivastava, M. B. Sensorsim: a simulation framework for sensor networks. *Proceedings of the the 3rd ACM international workshop on Modeling, analysis and simulation of wireless and mobile systems, Boston, MA, USA* (2000), 104–111.

[65] Petrovic, D., Shah, R. C., Ramchandran, K., and Rabaey, J. Data funneling: Routing with aggregation and compression for wireless sensor networks. *IEEE Sensor Network Protocols and Applications (SNPA)* (maio de 2003).

[66] Garey, M. R., and Johnson, D. S. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* Freeman, San Francisco, 1979.

[67] Madden, S., Szewczyk, R., Franklin, M. J., and Culler, D. Supporting aggregate queries over ad-hoc wireless sensor networks. *IEEE Workshop on Mobile Computing Systems and Applications (WMCSA)* (junho de 2002).

[68] Madden, S., Franklin, M. J., Hellerstein, J. M., and Hong, W. Tag: a tiny aggregation service for ad-hoc sensor networks. *Proceedings of the 5th symposium on Operating systems design and implementation* (2002).

[69] PC/104 CPU Modules. Página da World Wide Web. url: `http://www.adlogic-pc104.com/products/cpu/pc104/`, janeiro de 2004.

[70] MICA, MICA2 Motes and Sensors. Página da World Wide Web. url: `http://www.xbow.com/`, janeiro de 2004.

[71] Adjie-Winoto, W., Schwartz, E., Balakrishnan, H., and Lilley, J. The design and implementation of an intentional naming system. *Proceedings of the 17th Symposium on Operating Systems Principles* (dezembro de 1999), 186–201.

[72] Steinberg, A. N., Bowman, C. L., and White, F. E. Revisions to the JDL data fusion model. *Proceedings of the International Society for Optical Engineering (SPIE 1999), Orlando, FL, USA 3719* (abril de 1999), 430–441.

[73] Kalpakis, K., Dasgupta, K., and Namjoshi, P. Efficient algorithms for maximum lifetime data gathering and aggregation in wireless sensor networks. *The International Journal of Computer and Telecommunications Networking 42*, 6 (agosto de 2003), 697–716.

[74] Boulis, A., Ganeriwal, S., and Srivastava, M. Aggregation in sensor networks: An energy-accuracy trade-off. *IEEE Sensor Network Protocols and Applications (SNPA)* (maio de 2003).

[75] Hall, D. L. *Mathematical Techniques in Multi-Sensor Data Fusion.* Artech House, 1992.

[76] Klein, L. A. *Sensor and Data Fusion Concepts and Applications, Second Edition.* Tutorial Texts in Optical Engineering SPIE PRESS Vol. TT355, 1999.

[77] Jr., P. Z. P. *Probability, Random Variable and Random Signal Principles - Fouth Edition.* McGraw-Hill, Inc., 2001.

[78] Wu, H., Siegel, M., Stiefelhagen, R., and Yang, J. Sensor fusion using dempster-shafer theory. *IEEE Instrumentation and Measurement Technology Conference Anchorage* (maio de 2002).

[79] Stochero, J. M. Roteamento favorecendo agregação de dados em rede de sensores. *Dissertação de Mestrado. Universidade Federal do Rio de Janeiro, COPPE/GTA* (junho 2004).

[80] Heidemann, J., Silva, F., Yu, Y., Estrin, D., and Haldar, P. Diffusion filters as a flexible architecture for event notification in wireless sensor networks. *USC/ISI Technical Report 2002-556* (2002).

[81] Coffin, D., Hook, D. V., Govindan, R., Heidemann, J., and Silva, F. Network routing application programmer´s interface (API)and walk through 9.1. *Technical Report MIT/LL and USC/ISI* (junho de 2003).

[82] Fall, K., and Varadhan, K. The ns manual. the vint project. Página da World Wide Web. url: `http://www.isi.edu/nsnam/ns/ns-documentation.html`, janeiro de 2004.