

UMA ABORDAGEM DE FUSÃO DE DADOS COM RESTRIÇÕES DE TEMPO REAL EM REDES DE SENSORES SEM FIO

A. R. PINTO¹, BENEDITO R. BITENCORT², CARLOS MONTEZ^{1,2}

¹Programa de Pós-Graduação em Engenharia Elétrica – PGEEL/UFSC

²Programa de Pós-Graduação em Automação e Sistemas – PPGEAS/UFSC

Universidade Federal de Santa Catarina – Florianópolis – SC – 88040-900

E-mails: {arpinto,benedito,montez}@das.ufsc.br

Abstract: Wireless sensor networks are often deployed in order to detect signals generated by physical world. Data Fusion approaches can be used to increase decision dependability or decrease energy consumption. This paper consider a parallel data fusion approach with time constraints. Results comparing network density with fusion efficiency and messages average are showed. We demonstrate via simulations a strong relationship between the network density and data fusion efficiency. Moreover, an alternative approach is pointed out.

Keywords: Wireless Sensor Networks, Data Fusion, Real Time

Resumo: Redes de Sensores sem Fio são geralmente utilizadas para detecção de sinais gerado pelo mundo físico. Neste contexto, abordagens de fusão de dados são empregadas para aumentar a confiabilidade da decisão ou ainda diminuir gasto energético. Neste trabalho apresentamos uma abordagem de fusão de dados paralela com restrições temporais. Utilizamos testes de simulação para comparar o aumento de densidade da rede com a eficiência apresentada na fusão e número de mensagens/round. Resultados experimentais demonstram que a medida que a densidade da rede aumenta decai a eficiência da fusão. Além disso, apresentamos uma alternativa para abordagem proposta.

Palavras-chave: Redes de Sensores sem Fio, Fusão de Dados, Tempo Real

1. Introdução

Redes de sensores sem fio vêm sendo propostas como infraestrutura necessárias para as diversas novas aplicações de automação, envolvendo desde ambientes a céu aberto em sistemas petrolíferos até ambientes controlados de chão de fábrica em sistemas fabris [13].

O termo redes de sensores sem fio é usualmente empregado em muitas variações de composição e topologia. A rede pode ser composta por nodos homogêneos, ou, por outro lado, pode possuir nodos com diferentes capacidades ou com sensores diferenciados dos demais [1-3].

A questão energética é uma das características mais importantes dessas redes, pois o funcionamento dessas está intimamente ligado ao consumo otimizado das baterias dos seus nodos.

Muitas destas redes possuem alta densidade de nodos, são passíveis de falhas na comunicação e nos próprios nodos. Devido a estas características, a topologia destas redes é altamente dinâmica e imprevisível.

Objetivando aumentar a confiabilidade das leituras dos sensores, ou ainda, economizar energia dos nodos e aumentar a vida útil da rede, diversas abordagens utilizam técnicas de fusão de dados [6-10].

Em técnicas de fusão de dados, valores lidos dos sensores presentes nos nodos da rede podem ser enviados para uma estação base, a qual agrega esses

dados em busca de alguma informação específica da rede. Desta forma, não é mais necessário confiar na leitura de um único nodo, e sim, na combinação de diversas amostras. Por conseguinte, mesmo que haja alguns nodos faltosos, pode-se tomar decisões baseadas na rede como um todo. Esse é um dos paradigmas mais importantes desse tipo de redes: a confiabilidade da rede e dos dados monitorados não é baseada em nodos individuais, mas sim, no coletivo.

Apesar da fusão de dados ser baseada na leitura de sensores de diversos nodos, algumas abordagens de fusão de dados coletam amostras de apenas parte da rede, economizando a bateria de alguns nodos e prolongando a vida útil da rede [6,7,8,10].

Como essas redes são usadas para monitorar dados do mundo real, o conceito de “frescor” dos dados passa a ser importante. A validade da amostra coletada expira rapidamente. Alguns trabalhos atribuem valores de *deadlines* às mensagens para modelar essa restrição temporal explícita desse tipo de aplicação [3].

Este trabalho propõe uma abordagem de fusão paralela de dados [6] a qual considera que existem *deadlines* firmes (*firm deadlines*) associados às tarefas de envio de mensagens. Considera-se que a estação base recebe periodicamente os dados de sensores da rede, efetua fusão de dados e toma decisões baseadas nos resultados dessa fusão. Os *deadlines* das mensagens surgem, portanto, da necessidade da tomada de decisão dentro de um tempo máximo especificado.

Embora um dos compromissos da presente abordagem seja o de buscar o atendimento às restrições temporais (*deadlines*) das mensagens, a qualidade da fusão (mensurada aqui pela quantidade de amostras utilizadas para o processo de fusão de dados) e a questão energética são levadas em consideração.

Devido ao comportamento dinâmico das redes de sensores sem fio, não é possível assegurar o cumprimento dos *deadlines* das mensagens. No sentido de avaliar a abordagem, neste artigo é mostrada, através de simulações, a influência da densidade da rede de sensores sem fio, com relação às três métricas mencionadas: qualidade dos dados, atendimento dos *deadlines* e economia de energia (Figura 1).

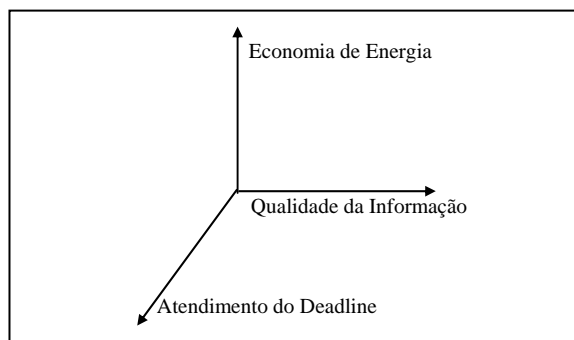


Figura 1: As três dimensões consideradas na abordagem.

Através dos dados obtidos pelas simulações mostra-se que é bastante promissor adotar-se uma abordagem probabilista, no envio de mensagens por cada nodo da rede.

Este artigo está organizado da seguinte forma: nas seções 2 e 3 apresentamos conceitos básicos de redes de sensores sem fio e descrevemos técnicas de fusão de dados. Na seção 4 são descritos o modelo e a abordagem proposta. Os resultados experimentais são apresentados na seção 5. Por fim apresentamos conclusões e direções futuras de pesquisa.

2. Redes de Sensores Sem Fio

Rede de Sensores Sem Fio é um termo geral que engloba diversas variações em composição e implantação. Uma rede de sensores típica consiste de um grande número de sensores implantados em um ambiente que será monitorado e, possivelmente, controlado. Geralmente, cada nodo da rede consiste de sensores, dispositivos de comunicação sem fio, unidade de processamento e bateria [3]. Um esquema mais detalhado é mostrado na Figura 2.

A restrição de recursos de software e hardware é um dos grandes desafios das redes de sensores sem fio. Cada nodo dessa rede, usualmente, possui restrições de energia e capacidade de memória e de processamento [1,2,3]. Algumas destas redes possuem alta densidade, os nodos são sujeitos a falhas e a topologia pode mudar dinamicamente.

Geralmente, uma estação base (*base station*) é empregada para coletar os dados obtidos pelos nodos sensores [3]. Estações base são nodos especiais (podendo ser mesmo um PDA ou computador) que geralmente possuem características especiais, como maior capacidade de armazenamento ou processamento.

Uma das principais mudanças de paradigma deve-se ao fato das redes de sensores serem centradas em dados. Ao contrário das redes de computadores convencionais – onde um usuário precisa conhecer o endereço de determinado equipamento para acessar seus recursos – nas redes de sensores o importante é o fenômeno a ser monitorado. Desta forma, pouco importa qual nodo está fornecendo a informação, e sim a localização do sinal coletado [2,3].

Algumas das características de uma rede de sensores sem fio são [2]:

- Comunicação baseada em difusão (*broadcast*) de pequeno alcance e roteamento *multi-hop*;
- Alta densidade e esforço cooperativo entre os nodos;
- Frequentes mudanças de topologia devido a falhas e desligamentos ou “sonos” voluntários (*sleeping*) nos nodos;
- Limitações de energia, potência de transmissão de dados, memória e poder de processamento.

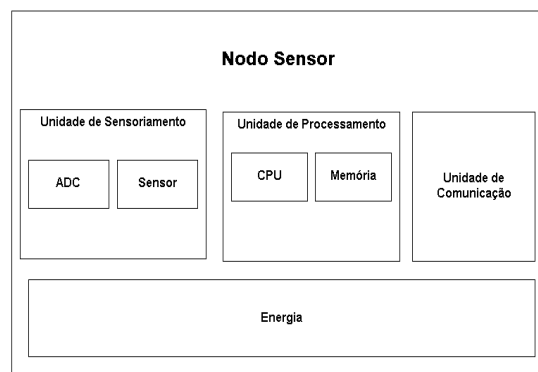


Figura 2. Esquema de arquitetura de um nodo sensor [1].

Essa tecnologia possui forte dependência e interação com o ambiente na qual está inserida. Devido a esta característica, e aliada ao fato de que as grandezas monitoradas rapidamente deixam de serem válidas, frequentemente restrições temporais são impostas às aplicações que envolvem essas redes [3]. Os dados coletados possuem um prazo de validade – essa característica é chamada de “frescor dos dados” (*data freshness*) [3].

A alta densidade de nodos que as redes de sensores sem fio podem apresentar, aliadas ao não-determinismo e grande incidência de ruídos devido ao meio de comunicação sem fio, são os maiores

desafios enfrentados por aplicações que possuam restrições de tempo real.

Aplicações para esta emergente tecnologia são diversas. Automação industrial, aplicações petrolíferas, casas inteligentes (domótica), monitoração ambiental e agrícola e sistemas de segurança, são apenas algumas das aplicações que podem utilizar-se das facilidades oferecidas pelas redes de sensores sem fio [1-5].

3. Fusão de Dados em Redes de Sensores Sem Fio

Segundo [8], existem duas formas básicas de troca de informações entre os dados. Quando dois nodos coletam dados correlacionados entre si, o envio de um pacote que colete cada leitura é necessário para um desempenho ótimo. Esta troca de informações é chamada de **fusão de dados**. Por outro lado, se as leituras são independentes, a **fusão de decisão** é suficiente. Neste último caso, somente valores de decisão (probabilísticos) são trocados. Embora esta nomenclatura possa ser usada, neste artigo utilizaremos sempre o termo “fusão de dados” para ambos os casos.

O principal objetivo de uma fusão de dados em redes de sensores sem fio é detectar um determinado evento ou coletar uma informação específica, ainda que na presença de ruídos. A detecção tipicamente é executada em um contexto colaborativo, combinando a leitura de múltiplos sensores [6].

A detecção colaborativa pode ser feita de forma paralela ou serial, levando em consideração a comunicação entre os nodos. A fusão de dados paralela, apresentada na Figura 3, é executada a partir de um número específico de nodos enviando suas leituras de sinal para uma estação base. A fusão serial, por sua vez, é combinada com técnicas de roteamento (Figura 4) onde as leituras são coletadas ao longo da rede [6].

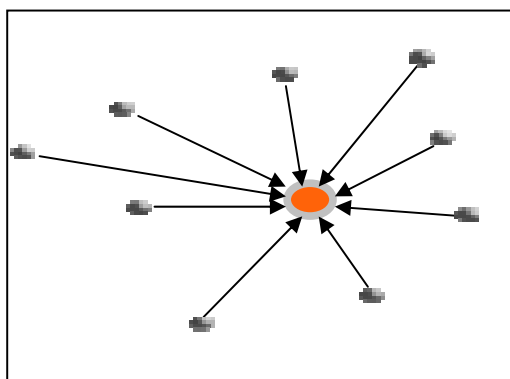


Figura 3. Fusão de dados paralela.

Um terceiro modo de comunicação entre os nodos envolvidos na fusão de dados seria a híbrida. Nesse tipo de fusão, a rede se estrutura na forma de uma árvore computacional com partes da rede sendo organizadas em agregados (*clusters* de nodos). Como

pode ser visto na Figura 5, esta abordagem pode ser considerada uma combinação entre a fusão de dados paralela e serial.

Devido à pouca previsibilidade das redes de sensores sem fio, a fusão de dados de vários sensores aumenta a precisão da detecção de determinado evento. Desta forma, não é preciso confiar na decisão de um único nó sensor, mas sim na leitura de diversos nodos da rede [6-9].

Além disso, a fusão de dados pode aumentar o tempo de vida da rede. Algumas abordagens visam diminuir o gasto energético. Segundo [6], em uma rede com alta densidade de nodos seria necessário coletar dados somente de alguns sensores para tomar determinada decisão. Dessa forma, um menor número de nodos executariam o sensoriamento e utilizariam a unidade de comunicação. O trabalho em [10] propõe a integração da fusão de dados na camada de rede, com o intuito de economizar energia.

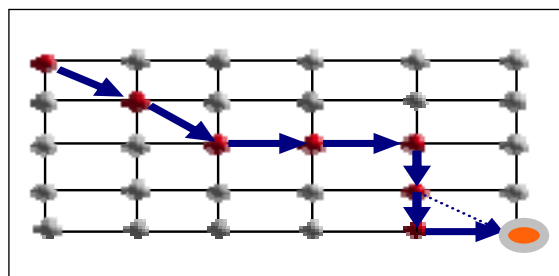


Figura 4. Fusão de dados serial.

Uma vez que as redes de sensores atuam no mundo real, requisitos de tempo real são impostos a este novo paradigma [3]. Além disso, há a necessidade de se considerar a confiabilidade da tomada de decisão, baseada na quantidade de sensores consultados. Neste trabalho denominamos este requisito: **qualidade da fusão**.

Portanto, a qualidade da fusão de dados, consumo energético e respeito aos *deadlines* da aplicação são qualidades importantes de uma aplicação de fusão de dados tempo real em redes de sensores sem fio (vide Figura 1).

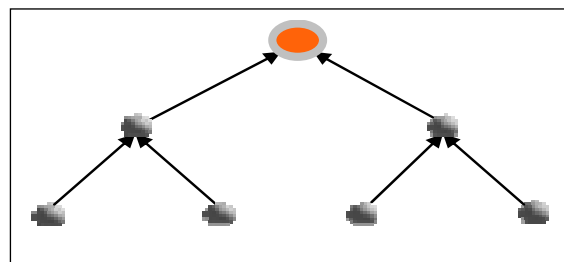


Figura 5. Fusão de dados híbrida.

É necessário estabelecer uma solução de compromisso entre: (i) aumentar o tempo de vida da rede, (ii) obter uma melhor qualidade de fusão de dados e (iii) atender os requisitos temporais da aplica-

ção. Essa necessidade de balancear o atendimento a esses requisitos se deve ao fato que a maximização de qualidade da fusão de dados, lendo dados de mais sensores, por exemplo, pode acarretar em um maior gasto energético. Da mesma forma, o atendimento de um *deadline* na tomada de decisão pode implicar na necessidade de reduzir o número de amostras coletadas nos nodos.

4. Um modelo de sistema e uma abordagem para fusão de dados

O modelo de sistema proposto considera a existência de duas entidades básicas: Mestre e Escravo. Cada rede possui um mestre e vários escravos. O mestre – que pode ser uma estação base – é responsável por receber as amostras coletadas pelos escravos e realizar a fusão de dados.

O nodo mestre possui duas tarefas periódicas: TRecebe e TFusão e uma fila de amostras coletadas dos nodos escravos. Ambas as tarefas são escalonadas por prioridades fixas (algoritmo taxa monotônica – *rate monotonic*) e são garantidas por um teste de escalonabilidade baseada em utilização.

TRecebe coleta as amostras enviadas pelos nodos escravos e armazena na fila de amostras. A tarefa TFusão calcula periodicamente a média das amostras, e caso as mensagens de todos os nodos não tiverem sido recebidas considera-se que a qualidade da fusão é afetada

Por exemplo, se ao término de TFusão houver somente três amostras em uma rede de cinco nodos, cumpriu-se o *deadline* da fusão de dados; porém, a qualidade (confiabilidade) da fusão de dados decaiu.

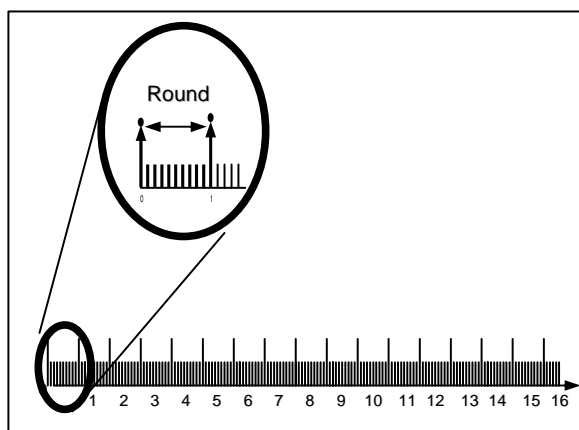


Figura 6. Modelo de tempo baseado em *rounds*.

O modelo de tempo real utilizado no modelo de sistema proposto é baseado no conceito de *round*. Dessa forma, a cada intervalo de tempo incrementa-se o número de *rounds* tanto no mestre quanto nos escravos. Logo, para que uma mensagem seja aceita pelo mestre para fusão, esta deve estar no *round* corrente. A Figura 6 representa o conceito de *round* utilizado.

Considera-se, portanto, que há sincronização de relógios na rede de sensores sem fio. Essa sincronização ocorre, por exemplo, pela presença de dispositivos GPS nos nodos sensores, que já são comercializados [11].

A presente abordagem não considera roteamento. Por conseguinte, todos os sensores estão na mesma área de alcance e podem se comunicar com um único salto (*hop*). Desta forma, utilizamos fusão de dados paralela com restrições temporais. Por simplicidade, mas sem prejudicar a generalidade, considera-se que a fusão de dados é feita com base em uma média simples:

$$f(x) = \frac{\sum_{i=1}^n a(i)}{N}$$

sendo que $a(i)$ é a amostra do sensor de identidade i e N é o número de amostras coletadas antes ao término da fusão de dados.

Consideramos como uma perda de *deadline* quando uma mensagem de um nodo não chega a tempo do processo de fusão de dados. Por exemplo, se há cinco nodos na rede e a fusão é feita com três amostras, são consideradas duas perdas de *deadline*. A qualidade da fusão é medida de acordo com a quantidade N , desta forma 100% de qualidade seria uma fusão feita com amostras de todos os nodos da rede.

5. Resultados Experimentais

Vários experimentos foram efetuados, através de simulação, considerando-se o modelo de sistema proposto com diversas densidades de nodos. Utilizou-se as densidades 1, 2, 5, 10, 15, 20, 25, 50, 75, 100, 150 e 200. As tarefas de envio dos nodos escravos tiveram seus *deadlines* fixados em 1 segundo (o mesmo valor de um *round*).

O simulador adotado foi o TrueTime [12], que possui nativamente facilidades para implementar sistemas com restrições temporais. O protocolo de comunicação simulado foi o IEEE 802.15.4 (ZigBee) com os parâmetros de rede padrões e tempo de simulação de 2000 segundos. O ZigBee vem sendo amplamente adotado em redes de sensores sem fio, e é a mesma tecnologia utilizada nos nodos de sensores Telos [11].

O objetivo principal desses experimentos foi observar a influência do crescimento densidade no atendimento dos *deadlines* e qualidade da informação da fusão de dados no nodo mestre. A idéia por trás da métrica “qualidade da informação” é a de mensurar a qualidade da leitura dos dados pela combinação de valores de diversos sensores. Desta forma, já que não é possível garantir a confiabilidade de poucos sensores quanto maior o número de leituras maior a qualidade da fusão de dados.

A Tabela 1 apresenta o total de mensagens recebidas pelo nodo mestre na simulação, média de men-

sagens/round e eficiência (porcentagem de mensagens recebidas).

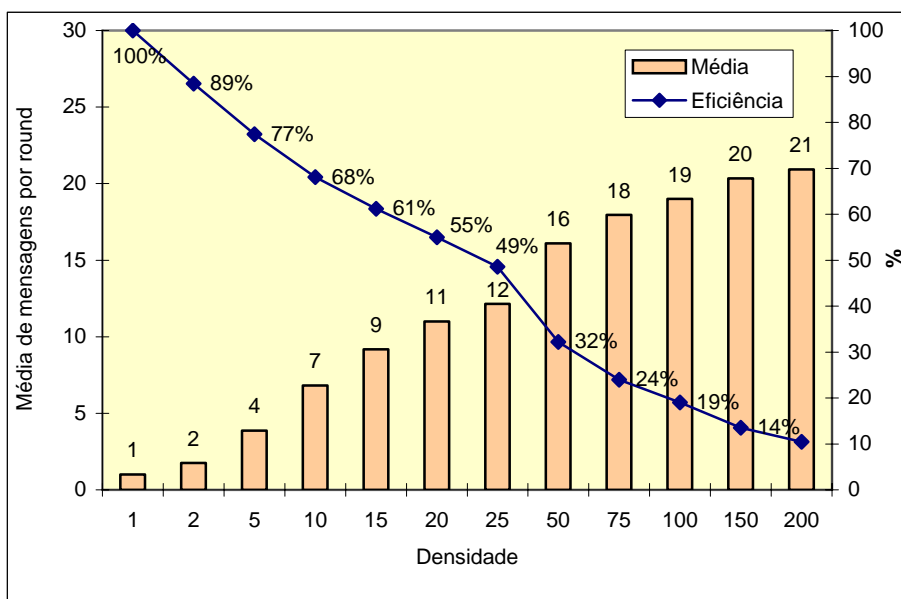


Figura 7. Relação entre mensagens recebidas por round e eficiência.

A Figura 7, baseada nos resultados da Tabela 1, apresenta a relação da média de mensagens/round com a eficiência.

Tabela 1: Resultados Obtidos

Densidade	Total de mensagens	Média	Eficiência (%)
1	2.000	1	100
2	3.528	1,76	88,5
5	7.744	3,87	77,4
10	13.614	6,80	68,1
15	18.536	9,17	61,2
20	21.990	10,99	55,0
25	24.282	12,14	49,0
50	32.210	16,10	32,2
75	35.936	17,96	24,0
100	38.022	19,01	19,0
150	40.665	20,33	13,6
200	41.847	20,92	10,5
Probabilista (densidade 20)	15.188	7,59	69,4

Analisando-se o gráfico da Figura 7 é possível notar que à medida que a densidade da rede aumenta, a eficiência decai e a média cresce lentamente. Logo, embora o número de nodos escravo aumente não é possível aproveitar todas as leituras do sensoriamento a tempo de realizar a fusão de dados no nodo mestre.

Portanto, essa métrica “eficiência” mensura indiretamente o consumo de energia. Quanto maior a

eficiência, menor o consumo de energia, pois despendesse menos energia para fazer com que as mensagens alcancem o mestre para sua fusão de dados.

A Figura 8 apresenta o gráfico de número total de mensagens recebidas que foram utilizadas na fusão de dados durante a simulação. Observa-se que, apesar da redução da eficiência, para o intervalo de densidade simulado o total de mensagens recebidas sempre cresce.

No entanto, uma vez que o maior gasto energético é com a comunicação, uma maior densidade de redes implica em um maior gasto energético. Por conseguinte, consideramos que o custo energético não compensa a pequena melhora da média de mensagens por round apresentada.

Resultados com abordagem probabilista

Através dos resultados obtidos na simulação observa-se que, à medida que a densidade aumenta, não é possível aproveitar todas as mensagens enviadas pelos nodos escravos. Desta forma, implementamos uma abordagem probabilista para comparar com os resultados obtidos.

Na abordagem probabilista proposta aqui, cada nodo sensor tem uma probabilidade x de enviar a mensagem a cada round. Essa probabilidade x é escolhida de acordo com a eficiência apresentada na Tabela 1. Por exemplo, um teste foi efetuado com densidade de 20 sensores e a probabilidade de envio foi de 55%.

Os resultados obtidos indicaram uma melhora significativa na eficiência (economia de energia) uma vez que a abordagem probabilista apresentou 14,5% de aumento de eficiência frente à primeira abordagem.

Apesar de todos os resultados obtidos dependerem do protocolo adotado (no caso, o ZigBee), acredita-se que outros protocolos de rede apresentem comportamento semelhantes devido às ocorrências de colisão e adoção de períodos de *backoff* entre as retransmissões.

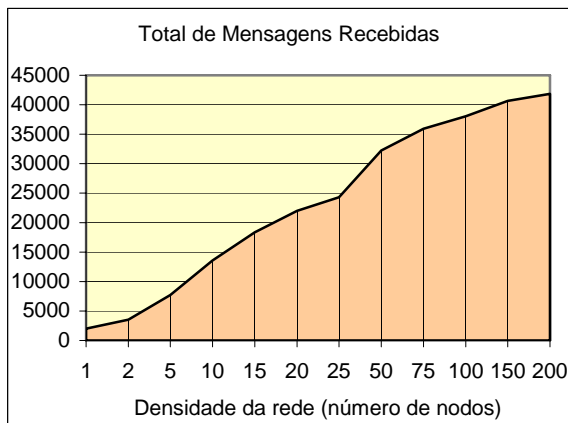


Figura 8. Total de mensagens recebidas pelo mestre com tempo de simulação de 2000 segundos.

6. Conclusão e Trabalhos Futuros

Uma rede de sensores sem fio com grande número de nodos quase sempre é interessante sob o ponto de vista que esta, se bem configurada, consegue se manter operacional por mais tempo que uma rede menor (através de técnicas de rodízio, escalonamento de sono, etc.). No entanto, mostrou-se neste trabalho que em redes com restrições temporais, muitas mensagens chegam após o *deadline* e pode ocorrer uma grande ineficiência nos recursos despendidos na transmissão.

Os resultados obtidos neste trabalho indicaram que à medida que a densidade da rede aumenta, menor é a eficiência da fusão paralela. Desta forma, esse tipo de abordagem acarreta em um maior gasto energético para que seja possível garantir certo nível de qualidade da fusão de dados.

No entanto, uma direção de pesquisa interessante apontada foi a de uma abordagem probabilista. Nesse caso, é possível trabalhar com uma alta densidade de rede e ainda assim aumentar o tempo de vida da rede e a qualidade da informação. Os resultados experimentais apresentaram uma melhora significativa na eficiência quando utilizamos a abordagem probabilista.

Como trabalhos futuros, indicamos a necessidade de aprimorar a abordagem probabilista apresentada. Além disso, testes com outros protocolos de comunicação e abordagens de fusão de dados serial e híbrida poderiam ser consideradas.

Agradecimentos

Este trabalho foi parcialmente apoiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

Referências Bibliográficas

- [1] Akyildiz, I.F., Su W., Sankarasubramaniam Y. e Cayirci E. (2002). A Survey on Sensor Networks. IEEE Communications Magazine, pp. 102-114.
- [2] Ilyas, M. e Mahgoub, I. Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems, 2004, CRC Press Inc., Boca Raton, Florida USA.
- [3] Stankovic, J.A., Abdelzaher, T.F., Lu, C., Sha, L., Hou, J.C., Real-Time Communication and Coordination in Embedded Sensor Networks. Proceedings of The IEEE, Vol. 91, No. 7, (Jul. 2003), 1002-1022.
- [4] Mainwaring, A., Culler, D., Polastre, J., Szewczyk, R., Anderson, J., Wireless Sensor Networks for Habitat Monitoring, Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications, 2002, 88-97.
- [5] Werner-Allen, G., Johnson, J., Ruiz, M., Lees, J., Welsh, M., Monitoring Volcanic Eruptions with a wireless Sensor Network, Proceedings of the Second European Workshop on Wireless Sensor Networks, 2005, 108-120.
- [6] Patil, S., Das, S., Nasipuri, A., Serial Data Fusion Using Space-filling Curves in Wireless Sensor Networks, IEEE SECON 2004, 2004, 182-190.
- [7] Pai, H., Han, Y., Power-Efficient Data Fusion Assurance Using Direct Voting Mechanism in Wireless Sensor Networks, IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing, 2006, 368-375.
- [8] Costa, A., Sayeed, A., Data Versus Decision Fusion for Distributed Classification in Sensor Networks, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003, 832-835.
- [9] Liu, E., Moura, J., Fusion in Sensor Networks: Convergence Study, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004, 865-870.
- [10] Liang, B., Liu Q. A Data Fusion Approach for Power Saving in Wireless Sensor Networks, First International Multi-Symposiums on Computer and Computational Sciences, IMSCCS '06, 2006, 582-586.
- [11] Crossbow Inc., <http://www.xbow.com>, Disponível em 31 de julho de 2007.
- [12] TrueTime Simulation of Networked and Embedded Control System, <http://www.control.lth.se/truetime/>. Disponível em 16 de abril de 2007.
- [13] Marrón, P., Minder, D., Embedded WiSeNts Research Roadmap, Information Society Technologies, IST-004400, November 2006.