Aspectos Práticos sobre o Consenso Bizantino entre Participantes Desconhecidos

Eduardo Adilio Pelinson Alchieri[1], Luiz Renato Tomelin[1], Alysson Neves Bessani[2], Joni da Silva Fraga[1]

[1]DAS, Universidade Federal de Santa Catarina, Florianópolis - Brasil
[2]LaSIGE, Universidade de Lisboa, Lisboa - Portugal

Abstract. *Agreement protocols form the basis for the solution of most problems involving reliable distributed systems. Although the consensus has been widely studied in classic environments, few studies have considered this problem in dynamic and self-organizing systems, where the set of participants are previously unknown. Recently, a solution to Byzantine fault-tolerant consensus with unknown participants (BFT-CUP) was proposed, which identifies the degree of knowledge, about the system composition, necessary and sufficient to solve the consensus. This paper complements these theoretical works by analyzing practical aspects about the BFT-CUP. First, this problem is analyzed in Mobile Ad Hoc Networks through simulations of several realistic scenarios, where we can identify what are the parameters and settings required to solve BFT-CUP with a high probability. Moreover, this paper studies the use of BFT-CUP protocols in Vehicular Ad Hoc Networks, by analyzing the vehicular connectivity of the city of Porto - Portugal.*

Resumo. *Protocolos de acordo formam a base para a solução da maioria dos problemas que envolvem sistemas distribuídos e confiáveis. Apesar de o consenso ter sido amplamente estudado em ambientes clássicos, onde o conjunto de participantes é conhecido, poucos trabalhos consideram este problema em ambientes dinâmicos e auto-organizáveis, onde os participantes da computação são, a priori, desconhecidos. Recentemente foi proposta uma solução para o problema do consenso Bizantino entre participantes desconhecidos, chamada BFT-CUP (Byzantine Consensus with Unknown Participants), a qual busca definir o grau de conhecimento, sobre a composição do sistema, necessário e suficiente para que este problema admita solução. Este trabalho complementa os resultados teóricos obtidos até então e busca analisar aspectos práticos da realização do BFT-CUP. Primeiramente, este problema é analisado em redes MANETs (Mobile Ad Hoc Networks) através de simulações de vários cenários realistas, onde é possível verificar quais são os parâmetros e configurações necessários para que os participantes do sistema consigam resolver o BFT-CUP com uma alta probabilidade. Além disso, também é apresentado um estudo acerca da utilização do BFT-CUP em redes VANETs (Veicular Ad Hoc Networks), através de uma analise sobre a conectividade veicular da cidade de Porto - Portugal.*

1. Introdução

O problema do consenso, Lamport *et al.* [Lamport et al. 1982], e de um modo geral os algoritmos de acordo, formam a base para a solução da maioria dos problemas encontrados no desenvolvimento de sistemas distribuídos confiáveis, pois possibilitam que os participantes da computação distribuída coordenem suas ações de forma a manter a consistência em seus estados e garantir o progresso do sistema. Este problema foi vastamente estudado em redes clássicas, onde o conjunto de processos que participam de determinada computação é estático e conhecido

por todos os participantes do sistema. Mesmo nestes ambientes, o problema de consenso não tem solução determinista na presença de uma única falha, considerando que as entidades se comportam de forma assíncrona [Fischer et al. 1985].

Considerando sistemas dinâmicos e auto-organizáveis, como redes MANETs (*Mobile Ad Hoc Networks*) e VANETs (*Veicular Ad Hoc Networks*), as dificuldades encontradas na elaboração de protocolos para resolver o consenso aumentam. Um conhecimento inicial sobre a composição do sistema é uma premissa que não pode ser adotada nestes ambientes. Assim, os participantes da computação (e os seus conhecimentos sobre a composição do sistema) não podem ser previamente determinados.

Neste sentido, Cavin *et al.* [Cavin et al. 2004, Cavin et al. 2005] realizaram os primeiros esforços para resolver o consenso em redes desconhecidas (ou sistemas dinâmicos e auto-organizáveis), definindo quais são as condições necessárias e suficientes para resolver o FT-CUP (*fault-tolerant consensus with unknown participants*), considerando o conhecimento sobre a composição do sistema e os requerimentos de sincronia (encapsulados em um detector de faltas). No trabalho apresentado em [Cavin et al. 2005], prova-se que para resolver o FT-CUP em um cenário com o mínimo de conhecimento sobre a composição do sistema é necessário que o mesmo seja síncrono. Greve *et al.* [Greve and Tixeuil 2007] comprovam que realmente existe uma relação entre o grau de conhecimento obtido pelos participantes e os requisitos de sincronia, apresentando uma solução para o FT-CUP em sistemas parcialmente assíncronos (mínimo de sincronia necessária para resolver o consenso em sistemas estáticos [Dwork et al. 1988], desconsiderando protocolos que empregam randomização [Aspnes 2003]), onde os participantes devem obter um maior conhecimento sobre a composição do sistema.

Recentemente, Alchieri *et al.* [Alchieri et al. 2008] estenderam este modelo e estudaram o consenso bizantino entre participantes desconhecidos, BFT-CUP (*Byzantine fault-tolerant consensus with unknown participants*), o qual define as condições necessárias e suficientes para resolver o consenso nestes ambientes, considerando que processos podem agir maliciosamente. Estas condições referem-se ao grau de conhecimento obtido pelos participantes e ao modelo de sincronia apresentado pelo sistema. As condições suficientes são representadas por um conjunto de algoritmos capaz de resolver o BFT-CUP dado que o sistema apresente as condições identificadas como necessárias para que o BFT-CUP admita solução.

Este trabalho apresenta uma abordagem prática sobre a realização do BFT-CUP [Alchieri et al. 2008]. Primeiramente, o comportamento dos protocolos que resolvem o BFT-CUP [Alchieri et al. 2008] é avaliado em uma rede MANET de topologia arbitrária, onde o grau de conhecimento teoricamente necessário para resolver este problema nem sempre é obtido pelos participantes. Nossas avaliações são baseadas em simulações onde os parâmetros tanto do BFT-CUP quando da rede MANET são variados. Além disso, este trabalho também apresenta um estudo acerca da possibilidade de utilização do BFT-CUP em redes VANETs, através de uma análise sobre a conectividade veicular apresentada pela cidade de Porto-Portugal.

O restante deste trabalho está organizado da seguinte forma. A Seção 2 apresenta os principais conceitos relacionados com o problema do consenso bizantino entre participantes desconhecidos (BFT-CUP) e descreve brevemente a solução para este problema proposta em [Alchieri et al. 2008]. O entendimento destes conceitos é fundamental para a compreensão das análises realizadas nas seções seguintes. As seções 3 e 4 analisam o comportamento do BFT-CUP em redes MANETs e VANETs, respectivamente. Por fim, a Seção 5 discute alguns trabalhos relacionados e a Seção 6 apresenta as conclusões deste trabalho.

2. O Problema do Consenso em Redes Desconhecidas

2.1. Modelo de Sistema

Considera-se um sistema distribuído composto por um conjunto finito Π de processos retirados do conjunto universo U. Em uma *rede conhecida*, Π é conhecido por todos os processos do sistema, enquanto que em uma *rede desconhecida*, um processo $i \in \Pi$ poderá conhecer apenas um subconjunto $\Pi_i \subseteq \Pi$ (visão parcial do sistema).

Os processos estão sujeitos a faltas Bizantinas [Lamport et al. 1982], onde o limite de falhas f é conhecido. Comunicação entre processos conhecidos dá-se através de canais ponto a ponto confiáveis e autenticados, onde é necessário um protocolo de roteamento tolerante a faltas bizantinas [Kotzanikolaou et al. 2005]. Já a comunicação entre processos desconhecidos é realizada através de redundância de mensagens [Alchieri et al. 2008].

Além disso, os protocolos do BFT-CUP são projetados para sistemas assíncronos mas utilizam um camada subjacente de consenso bizantino clássico que pode ser implementada sobre sistemas parcialmente síncronos [Dwork et al. 1988] (ex., Paxos Bizantino [Castro and Liskov 2002]) ou sobre sistemas completamente assíncronos (ex., *randomized consensus* [Correia et al. 2006, Bracha 1984, Ben-Or 1983]). Desta forma, o BFT-CUP exige o mesmo nível de sincronia do protocolo de consenso clássico utilizado. Nas simulações realizadas neste trabalho, utilizamos o protocolo Paxos Bizantino [Castro and Liskov 2002], considerando um sistema parcialmente síncrono.

2.2. Definição

O problema do consenso consiste em garantir que todos os processos corretos acabarão por decidir pelo mesmo valor, que foi previamente proposto por algum processo. Formalmente, o consenso é definido pelas seguintes propriedades [Chandra and Toueg 1996, Castro and Liskov 2002]:

- *Validade*: se um processo correto decide pelo valor v, então v foi proposto por algum processo;
- *Acordo*: dois processos corretos não decidem por valores diferentes;
- *Terminação*: todo processo correto acaba por decidir por algum valor;
- *Integridade*: todo processo correto decide uma única vez.

2.3. Detectores de Participação

Para resolver qualquer problema distribuído não trivial, é necessário que os processos obtenham um conhecimento ao menos parcial sobre a composição do sistema caso cooperação entre os mesmos seja necessária. Os detectores de participação (PD) são oráculos distribuídos que fornecem informações sobre participantes do sistema. Cada participante i possui o seu detector de participação $i.PD$ e uma consulta a este detector retorna um subconjunto de processos de Π com os quais i pode colaborar [Cavin et al. 2004].

Detectores de participação fornecem um contexto inicial sobre a composição do sistema através do qual é possível expandir o conhecimento sobre Π. Assim, esta abstração enriquece o sistema com um grafo de conectividade por conhecimento. Este grafo é orientado, pois o conhecimento fornecido pelos detectores de participação não é necessariamente bidirecional [Cavin et al. 2004].

Definição 1 ***Grafo de Conectividade por Conhecimento****: Seja $G_{di} = (V, \xi)$ um grafo orientado representando a relação de conhecimento determinada pelo oráculo PD. Então, $V = \Pi$ e $(i, j) \in \xi$ se e somente se $j \in i.PD$ (i conhece j).*

Definição 2 ***Grafo Não-Orientado de Conectividade por Conhecimento****: Seja $G = (V, \xi)$ um grafo não orientado representando a relação de conhecimento determinada pelo oráculo PD. Então, $V = \Pi$ e $(i, j) \in \xi$ se e somente se $j \in i.PD \vee i \in j.PD$ (i conhece j ou j conhece i).*

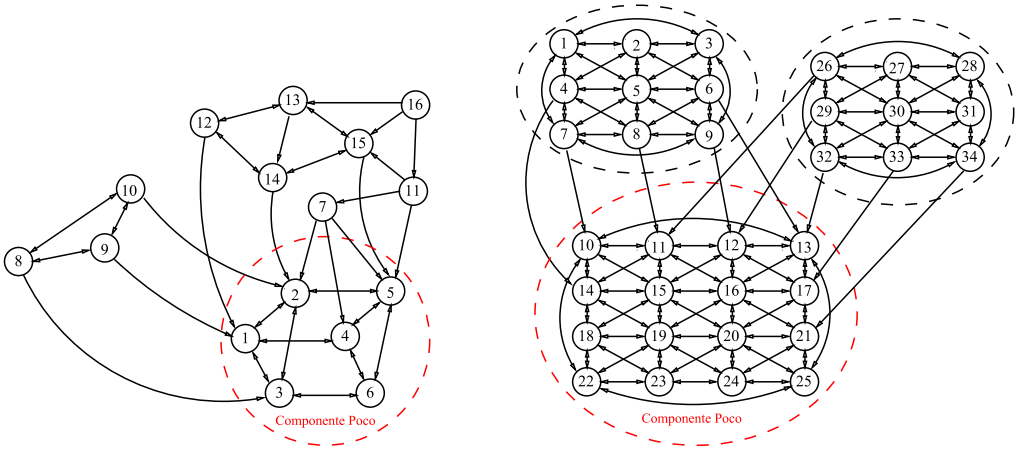
Baseado nas propriedades apresentadas pelos grafos de conectividade por conhecimento, algumas classes de detectores de participação foram propostas para resolver o consenso entre participantes desconhecidos. Neste sentido, o mínimo de conhecimento necessário para que o BFT-CUP admita solução é encapsulado em um PD k-OSR (k-One Sink Reducibility ou k-Redutível a Único Poço) [Greve and Tixeuil 2007, Alchieri et al. 2008].

Antes de definir como um detector de participação k-OSR encapsula o conhecimento sobre a composição do sistema, vamos definir algumas notações sobre grafos. Uma componente G_c de G_{di} é *k-fortemente conexa* se para qualquer par (i,j) de nós em G_c, i pode alcançar j através de pelo menos k caminhos disjuntos nos nós. Além disso, uma componente G_{sink} de G_{di} é uma *componente poço* quando não existe caminhos partindo de nós pertencentes à G_{sink} para outros nós de G_{di}, com exceção dos nós na própria componente G_{sink}.

Definição 3 ***PD k-Redutível a Único Poço (k-OSR)****: O grafo de conectividade por conhecimento G_{di}, que representa a relação de conhecimento induzida pelo oráculo PD, satisfaz as seguintes condições:*

1. *O grafo não orientado G obtido de G_{di} é conexo;*
2. *A redução de G_{di} as suas componentes fortemente conexas tem apenas uma componente poço, chamada G_{sink};*
3. *G_{sink} é k-fortemente conexa;*
4. *Para cada $i \notin G_{sink}$ e $j \in G_{sink}$, existem pelo menos k caminhos disjuntos nos nós de i para j.*

Para melhor ilustrar a Definição 3, a Figura 1 apresenta dois grafos G_{di} induzidos por um detector de participação k-OSR. As figuras 1(a) e 1(b) mostram relações de conhecimento induzidas por detectores de participação da classe 3-OSR e 5-OSR, respectivamente. Por exemplo, na Figura 1(a), o valor retornado por $1.PD$ é o subconjunto $\{2, 3, 4\}$ de Π.



(a) 3-OSR (b) 5-OSR

Figura 1. Grafos de Conectividade por Conhecimento.

2.4. BFT-CUP: Consenso Bizantino entre Participantes Desconhecidos

Esta seção apresenta os principais protocolos que formam a solução para o BFT-CUP [Alchieri et al. 2008], a qual é modular: primeiramente, um conjunto de participantes que compartilham a mesma visão parcial do sistema é identificado; após, um algoritmo de consenso para redes tradicionais (estáticas) é executado por estes participantes que difundem o valor de decisão pelo sistema, permitindo que todos os participantes decidam pelo mesmo valor.

Seguindo esta abordagem e utilizando os requisitos mínimos de sincronia para executar o algoritmo de consenso estático (clássico) Paxos Bizantino [Castro and Liskov 2002] (sistema parcialmente síncrono [Dwork et al. 1988]), é possível resolver o BFT-CUP utilizando um PD k-OSR, onde $k \geq 2f + 1$ e G_{sink} sendo composta por no mínimo $3f + 1$ participantes [Alchieri et al. 2008], que executam o algoritmo de consenso Paxos Bizantino.

Para analisar os aspectos práticos acerca do comportamento do BFT-CUP (seções 3 e 4), é fundamental entender o funcionamento de seus protocolos bem como das relações de dependência entre os mesmos. Desta forma, os algoritmos que solucionam o BFT-CUP são brevemente descritos a seguir. Os pseudocódigos destes algoritmos, bem como as provas de que os mesmos funcionam conforme especificado, podem ser encontrados em [Alchieri et al. 2008].

1 - *Reachable Reliable Broadcast*: Este protocolo é usado para comunicação entre participantes ainda não conhecidos, i.e., para comunicações realizadas antes dos participantes do sistema determinarem suas relações de conhecimento. Este protocolo basicamente realiza *flood* de mensagens no sistema, onde cada participante envia/encaminha as mensagens para todos os seus vizinhos em G_{di} (i.e., nós retornados pelo seu detector de participação). Todos os processos alcançáveis em G_{di} entregam estas mensagens assim que as autenticidades das mesmas sejam comprovadas através da redundância no recebimento, o que ocorre quando as mesmas forem recebidas através de pelo menos $f + 1$ caminhos disjuntos nos nós [Alchieri et al. 2008].

2 - *Discovery*: O primeiro passo para resolver o consenso entre participantes desconhecidos é fazer com que cada participante obtenha o máximo de conhecimento possível sobre a composição do sistema. Neste sentido, partindo do conhecimento inicial retornado pelos detectores de participação, este protocolo é usado pelos participantes do sistema para expandir seus conhecimentos sobre a composição do mesmo, através de uma espécie de busca em G_{di}.

Como os processos ainda não determinaram suas relações de conhecimento, este procedimento utiliza o algoritmo de difusão anteriormente descrito para garantir a integridade das mensagens recebidas pelos participantes [Alchieri et al. 2008]. Após a execução deste algoritmo, e por consequência da determinação das relações de conhecimento, os participantes comunicam-se através de canais ponto a ponto confiáveis e autenticados, onde um protocolo de roteamento tolerante a faltas bizantinas (ex.: [Kotzanikolaou et al. 2005]) deve ser empregado. Esta forma de comunicação também é utilizada pelo protocolo de difusão (*reachable reliable broadcast*) para enviar/encaminhar mensagens para nós vizinhos (que sempre são conhecidos pelo participante que está enviando/encaminhando uma mensagem).

Esta é a parte mais importante do BFT-CUP, onde cada participante obtém a sua visão parcial do sistema. Neste sentido, um dos grandes desafios do BFT-CUP é a busca de uma solução para o consenso Bizantino entre processos que possuem diferentes visões parciais sobre a composição do sistema.

3 - *Sink*: Após um participante obter a sua visão parcial do sistema, este algoritmo é utilizado para determinar se o mesmo pertence à componente G_{sink} de G_{di}. O princípio de funcionamento deste protocolo baseia-se no fato de que, através do algoritmo *Discovery*, cada participante em G_{sink} obtém a mesma visão parcial do sistema que é composta justamente pelos participantes de G_{sink}. Desta forma, este algoritmo isola um conjunto de participantes que compartilham a mesma visão do sistema, i.e., os participantes em G_{sink}.

4 - *Consensus*: Este algoritmo é executado após cada participante determinar se pertence à componente G_{sink} e apresenta comportamento distinto, dependendo da sua localização em G_{di}:

- Participantes em G_{sink}: Estes participantes executam um consenso bizantino clássico (ex. Paxos Bizantino [Castro and Liskov 2002]) e difundem o valor de decisão para os outros nós do sistema, que apenas solicitam este valor. Isso é possível graças ao fato dos participantes em G_{sink} compartilharem a mesma visão parcial do sistema. Este também é o motivo pelo qual G_{sink} deve possuir pelo menos $3f + 1$ participantes (o mínimo necessário para executar o consenso bizantino clássico [Castro and Liskov 2002]).
- Participantes que não pertencem à G_{sink}: Estes participantes apenas enviam uma requisição aos nós em G_{sink} solicitando o valor de decisão. Cada participante decide por um valor v quando receber pelo menos $f + 1$ respostas indicando que v foi o valor decidido pelos participantes em G_{sink}, garantindo a presença de pelo menos uma resposta produzida por um participante correto.

Relações de Dependência entre os protocolos do BFT-CUP. A Figura 2 apresenta as relações de dependência entre os protocolos utilizados para resolver o BFT-CUP. Note que o protocolo de difusão de mensagens (*reachable reliable broadcast*), juntamente com a camada subjacente de roteamento tolerante a faltas bizantinas, abstraem todos os aspectos relacionadas com a comunicação entre os processos. Um processo i deve invocar uma execução do consenso através do protocolo *consensus*, que utiliza o protocolo *sink* para verificar se i pertence à G_{sink}. Caso $i \in G_{sink}$, então i executa um consenso Bizantino clássico. De outra forma, i apenas solicita o valor de decisão conforme anteriormente descrito.

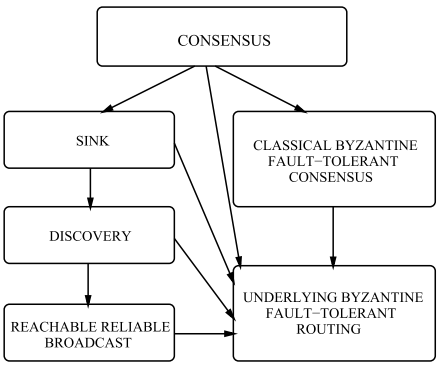


Figura 2. Relações de Dependência entre os Protocolos do BFT-CUP.

O protocolo *sink*, por sua vez, utiliza o protocolo *Discovery* para determinar a visão parcial que i obtém sobre o sistema. O algoritmo *Discovery* utiliza o conhecimento reportado pelo detector de participação de i ($i.PD$) para realizar uma busca em G_{di} obtendo o máximo de conhecimento possível sobre a composição do sistema. Para isso, também utiliza o algoritmo de difusão de mensagens, pois neste estágio os processos ainda não definiram completamente

as suas relações de conhecimento e i não pode enviar mensagens diretamente para todos os processos alcançáveis em G_{di} (nem todos estes processos já são conhecidos por i neste momento).

3. BFT-CUP em Redes MANETs

Esta seção busca analisar, através de uma série de simulações, o comportamento dos protocolos do BFT-CUP quando executados em MANETs. Para isso, tais protocolos foram implementados e simulados no *Jist/SWANS*. O *Jist* (*Java in Simulation Time*) [Barr et al. 2004] é um simulador de alta performance baseado em eventos discretos que executa sobre uma máquina virtual *Java*. Já o *SWANS* (*Scalable Wireless Ad hoc Network Simulator*) [Barr et al. 2005] é um módulo para simulação de MANETs construído sobre o *Jist*, sendo portanto também implementado em *Java*.

3.1. Implementação dos Detectores de Participação

Os detectores de participação são abstrações que fornecem conhecimento sobre a composição do sistema, sendo que para a solução do BFT-CUP o mínimo exigido é o detector de participação da classe k-OSR (Seção 2.3). No entanto, não existe uma proposta de implementação para um detector de participação que garante os atributos definidos para esta classe em uma rede MANET. A possibilidade de existência de tal implementação continua sendo uma questão em aberto [Costa et al. 2009]. Deste modo, em nossas simulações utilizamos o detector de participação **PD-1hop** [Costa et al. 2009] e buscamos avaliar a possibilidade de solução do BFT-CUP em um cenário onde o detector de participação não garante as propriedades do PD k-OSR. Note que caso o detector de participação assegure as propriedades do PD k-OSR, então o BFT-CUP sempre admite solução [Alchieri et al. 2008], i.e., a conectividade por conhecimento fornecida pelo detector é suficiente para resolver o BFT-CUP.

O detector de participação **PD-1hop** é muito simples e retorna para os processos uma lista contendo os nós que se encontram no seu alcance de transmissão (nós vizinhos). Sempre que este detector é iniciado em algum processo, o mesmo envia mensagens de "hello" através de um *broadcast* local, por meio da camada MAC da rede MANET, que atinge todos os seus vizinhos dentro de 1 salto. Ao receber estas mensagens, cada vizinho i adiciona a fonte (emissor) a sua lista de nós conhecidos, que é retornada em $i.PD$.

Esse procedimento de descoberta inicial termina quando um tempo é atingido. Durante este período, os processos podem difundir novas mensagens de "hello". Este detector é implementado de forma independente dos parâmetros do BFT-CUP e é ativado no início de cada simulação para retornar uma aproximação da vizinhança real na rede.

3.2. Configurações das Simulações

As configurações gerais utilizadas nas simulações foram adotadas com base nos trabalhos de Costa *et al.* [Costa and Greve 2007, Costa et al. 2009]. Estas configurações podem ser divididas em configurações da rede MANET (Tabela 1) e configurações do BFT-CUP (Tabela 2).

Os principais parâmetros da rede MANET variados nas simulações e que influenciam na possibilidade de solução do BFT-CUP são: *(1)* o alcance de transmissão e *(2)* a densidade de nós no sistema (quantidade de nós). Estes dois parâmetros estão diretamente relacionados com o grau de conhecimento que os participantes obtêm sobre a composição do sistema, i.e., com a conectividade por conhecimento apresentada em G_{di}, que é o fator determinante para a possibilidade de solução do BFT-CUP.

Os detectores de participação utilizados nestas simulações foram ativados durante $60s$ antes de cada execução do BFT-CUP. Durante este período, cada participante esperou um tempo

aleatório de até $20s$ para reenviar a mensagem de "hello" indicando sua presença no sistema (Seção 3.1). Vale ressaltar que estes parâmetros fizeram com que cada participante enviasse em média de 3 a 4 destas mensagens.

Parâmetros das Simulações (Simulador)	
Simulador	*Jist/SWANS*
Quantidade de nós	de 0 à 50
Área	$500X500\ m^2$
Alcance *wireless*	de $0m$ à $300m$
Repetições	10
Padrão de Mobilidade	*Randon WayPoint*
Velocidade dos nós	de $0m/s$ à $5m/s$
Tempo de pausa	de até $2s$

Tabela 1. Configurações da rede MANET.

Além disso, também vale destacar que apesar do sistema estar configurado para suportar um determinado número de faltas, que varia de acordo com as configurações utilizadas em determinada simulação (de 1 até 10 faltas), não foram introduzidos nós faltosos no sistema.

Parâmetros das Simulações (BFT-CUP)	
Tempo de PD	$60s$
Tempo para reenvio de "hello"	de até $20s$
Número máximo de faltas suportadas	de 1 à 10

Tabela 2. Configurações do BFT-CUP.

Por fim, o tempo de término das simulações foi definido como sendo o instante de tempo em que todos os eventos produzidos pelos nós foram processados pelo simulador, i.e., as simulações permaneceram em execução enquanto tinham eventos a serem processados.

3.3. Métricas Analisadas

Estas simulações têm por objetivo verificar em quais condições os participantes do sistema conseguem adquirir um conhecimento suficiente para resolver o BFT-CUP. Para atingir este objetivo, não é importante analisar o tempo necessário para a execução do BFT-CUP, mas sim as métricas diretamente relacionadas com a convergência do BFT-CUP. Neste sentido, as seguintes métricas serão analisadas nestas simulações:

1. *Conhecimento retornado pelos detectores de participação*: Esta métrica define a quantidade de conhecimento sobre a composição do sistema que os nós recebem diretamente dos detectores de participação. Deste modo, esta métrica refere-se ao conhecimento adquirido pelos participantes antes da execução do agoritmo *Discovery*.
2. *Convergência para Poço*: Esta métrica refere-se a quantidade de participantes que acabam por determinar que pertencem ao G_{sink} de G_{di}. No entanto, como veremos a seguir, em determinadas configurações é possível ocorrer execuções do BFT-CUP onde diferentes participantes acreditam que pertencem a diferentes componentes G_{sink}. Neste cenário é teoricamente possível ocorrer uma de duas coisas: *(1)* ou os participantes não terminam a execução do BFT-CUP (porque cada G_{sink} não possui nós suficientes para executar o consenso clássico); *(2)* ou ocorre desacordo nos valores das decisões (participantes em diferentes G_{sink} podem decidir por valores diferentes [Alchieri et al. 2008]).

3. *Terminação*: Esta métrica define a quantidade de participantes que terminam a execução dos protocolos do BFT-CUP. Rigorosamente falando, os nós em G_{sink} não podem terminar o BFT-CUP quando decidem por algum valor, pois os mesmos devem responder às requisições de outros nós solicitando este valor da decisão. Assim, tais nós devem aguardar por estas requisições enquanto o valor de decisão é importante para a aplicação que está utilizando o BFT-CUP como base para algum algoritmo distribuído. Nestas simulações, os nós em G_{sink} aguardaram por requisições de valor de decisão enquanto tinha algum evento para ser processado pelo simulador, i.e., enquanto existia a possibilidade de algum processo ter solicitado este valor de decisão.
4. *Acordo*: Esta métrica especifica a quantidade de participantes que terminam a execução dos protocolos do BFT-CUP decidindo pelo mesmo valor.

3.4. Resultados e Análises

Esta seção apresenta os resultados obtidos através das simulações, bem como as análises sobre estes resultados. Primeiramente, a Figura 3 apresenta os resultados para simulações onde fixou-se o número de faltas suportadas pelo sistema (f) em 1 falta e variou-se a densidade e o alcance de transmissão dos nós.

Na Figura 3(a) é possível observar o grau de conhecimento sobre a composição do sistema que é obtido através dos detectores de participação (conhecimento obtido antes da execução do algoritmo *Discovery*). Considerando a porcentagem de conhecimento inicial fornecido aos participantes (e não os números absolutos), é possível verificar que a porcentagem de conhecimento está diretamente relacionada com o alcance de transmissão dos nós, sendo que a densidade de nós presentes no sistema exerce pouca influência sobre este percentual. O conhecimento reportado pelos detectores de participação variou de aproximadamente 0% (para alcance de $30m$) a aproximadamente 50% (para alcance de $300m$). Considerando o cenário para alcance de $300m$, isto significa que quando o sistema foi composto por 4 nós, o conhecimento sobre a presença de mais 2 outros nós em média foi retornado pelos detectores de participação. Já para o cenário de 50 nós, este valor salta para 25 outros nós em média.

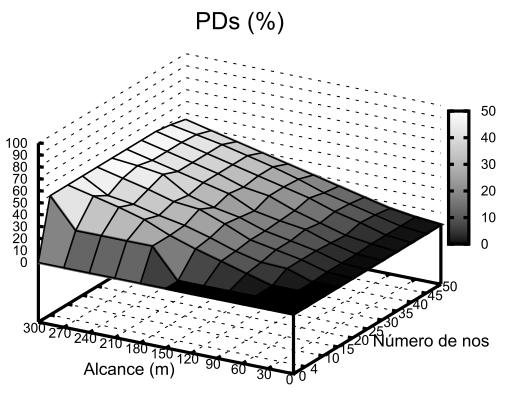
É importante notar que a união destes conhecimentos, obtidos em cada participante do sistema através de seus detectores de participação, gera o grafo de conectividade por conhecimento G_{di}, o qual indicará se a conectividade obtida é suficiente para resolver o BFT-CUP.

Após esta etapa de descoberta inicial, os nós executam o algoritmo *Discovery* para expandir este conhecimento e na sequência *Sink* para determinar se pertencem à G_{sink} de G_{di}. Neste sentido, a Figura 3(b) apresenta a porcentagem de participantes que acreditaram pertencer à G_{sink} nos mais variados cenários. Este gráfico apresenta três zonas distintas:

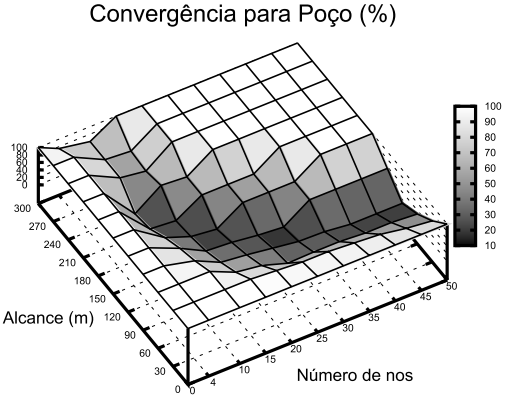
1. Alcance de transmissão e/ou densidade de nós baixo: Nestes cenários todos os nós acreditam pertencer a G_{sink}. No entanto, como os conhecimentos obtidos sobre a composição do sistema (depois do algoritmo *Discovery*) são muito reduzidos, os nós acreditam pertencer a diferentes componentes G_{sink}. Em alguns cenários os nós descobriram apenas outro ou até mesmo nenhum outro participante.
2. Alcance de transmissão e/ou densidade de nós moderada: Estes cenários formam o fundo da "vala" na Figura 3(b), onde o conhecimento obtido pelos nós é tão divergente entre os mesmos (embaralhado) que nenhum nó (ou poucos) acreditam pertencer a G_{sink}.
3. Alcance de transmissão e/ou densidade de nós alta: Nestes cenários todos os nós decidiram que pertencem ao mesmo G_{sink}, pois adquiriram um grande grau de conhecimento sobre a composição do sistema. Em praticamente todos estes cenários, os nós acabaram conhecendo todos os outros participantes do sistema.

As figuras 3(c) e 3(d) apresentam os gráficos para terminação e acordo, respectivamente. Como podemos perceber, estas métricas estão diretamente relacionadas com a convergência para um único G_{sink}, sendo que não é possível resolver o BFT-CUP nos cenários descritos anteriormente pelos itens 1 e 2. Nestes casos, o consenso tradicional não é executado pelos seguintes motivos:

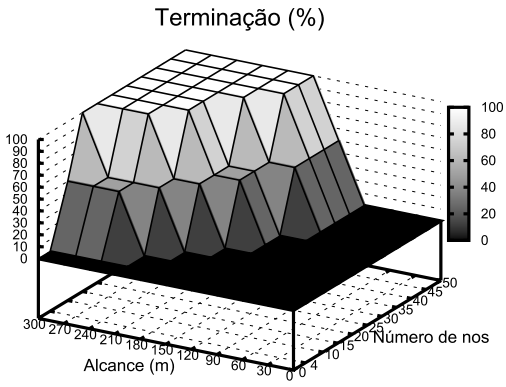
1. Cenários do item 1: Apesar de todos os nós acreditarem pertencer a G_{sink}, o consenso tradicional não é executado porque são diferentes componentes G_{sink} visto que os seus conhecimentos são muito reduzidos. De fato, nenhum nó consegue descobrir pelo menos outros 3 nós para executar o consenso clássico que exige $3f + 1$ participantes [Castro and Liskov 2002] (como $f = 1$, é necessário 4 participantes).
2. Cenários do item 2: O consenso clássico não é executado porque um número insuficiente de nós (menos de $3f + 1$ [Castro and Liskov 2002]) acreditam pertencer à G_{sink}.



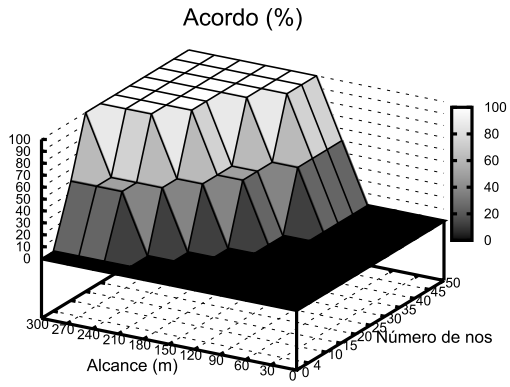
(a) Conhecimento adquirido através dos PDs.



(b) Convergência para Poço.



(c) Terminação.

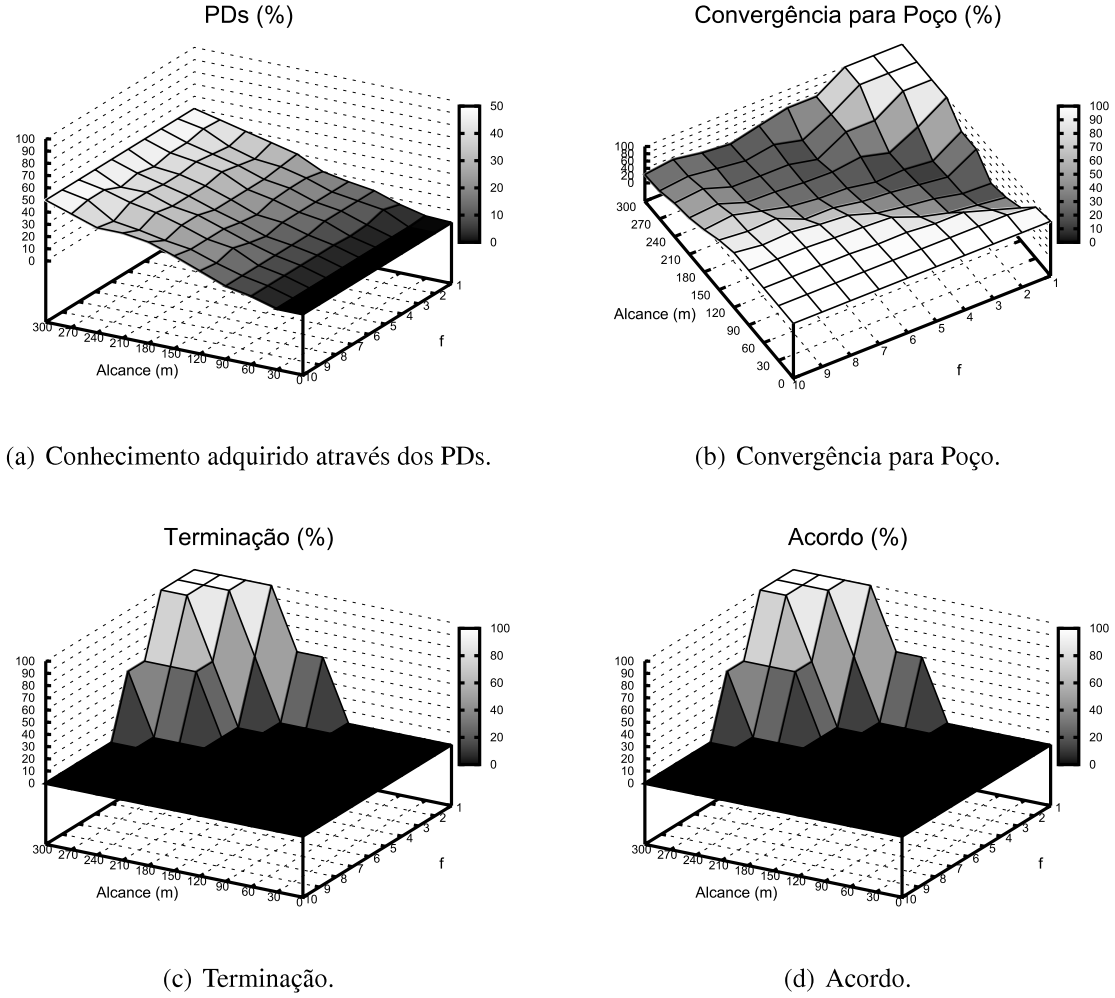


(d) Acordo.

Figura 3. Resultados das simulações para $f = 1$. Tanto o aumento da densidade de nós no sistema, quanto o aumento do alcance de transmissão dos nós, faz com que os participantes obtenham um melhor conhecimento sobre a composição do sistema, aumentando a conectividade em G_{di} e com isso também aumentando a probabilidade do BFT-CUP admitir solução.

Por estes motivos, os gráficos de acordo e terminação são idênticos representando que quando os algoritmos do BFT-CUP terminaram, a propriedade de acordo do consenso foi atendida. Este é um resultado importante, pois nestas simulações sempre que os nós terminaram a execução dos protocolos acabaram decidindo pelo mesmo valor.

A fim de analisar o comportamento do BFT-CUP em relação ao número de faltas suportadas pelo sistema, outros cenários foram simulados para o sistema composto por 30 nós. Os resultados são apresentados na Figura 4, onde é possível perceber que o número de faltas suportadas não têm relação com a quantidade de conhecimento obtida através dos detectores de participação 4(a). De fato, os detectores de participação trabalham de forma independente do BFT-CUP e este parâmetro não deve interferir em seu funcionamento.



(a) Conhecimento adquirido através dos PDs.

(b) Convergência para Poço.

(c) Terminação.

(d) Acordo.

Figura 4. Resultados das simulações para 30 nós. A probabilidade do BFT-CUP admitir solução em uma rede MANET é diretamente proporcional ao alcance de transmissão dos nós e inversamente proporcional ao número de faltas toleradas pelo sistema.

Nestas simulações foi possível resolver o BFT-CUP em cenários para f de até 5 faltas e alcance de transmissão a partir de $150m$. O BFT-CUP nunca terminou nos cenários com f maior do que 5 faltas, pois nestes casos não foi possível formar um único G_{sink}. A Figura 4(b) apresenta a porcentagem de nós que acreditaram pertencer a G_{sink}, cuja interpretação é da mesma forma que o anterior (Figura 3(b)): *(1)* alcance de transmissão baixo – os nós acreditaram que pertenciam à diferentes componentes G_{sink}; *(2)* alcance de transmissão moderado e/ou f alto – poucos nós acreditaram pertencer à G_{sink};*(3)* alcance de transmissão alto e/ou f baixo – os nós decidiram que pertenciam à mesma componente G_{sink}.

Novamente, sempre que os protocolos do BFT-CUP terminaram (Figura 4(c)) os nós acabaram decidindo pelo mesmo valor (Figura 4(d)), indicando que não ocorreu desacordo nos cenários simulados.

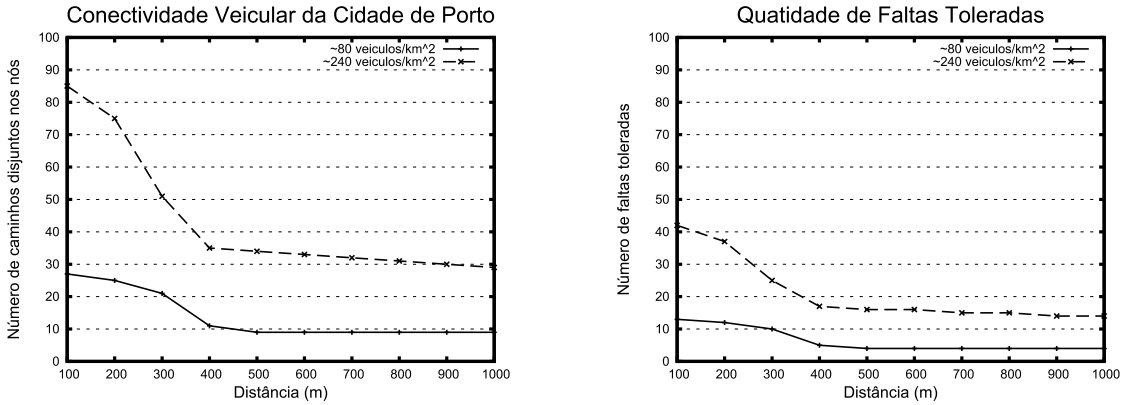
4. BFT-CUP em Redes VANETs: Exemplo Prático da Cidade de Porto - Portugal

Esta seção apresenta uma análise acerca da realização do BFT-CUP entre veículos (ou uma parte deles) que trafegam nas vias de uma cidade, ou seja, em uma rede VANET. O cenário escolhido foi a cidade de Porto - Portugal, cujo tráfego veicular foi modelado por Ferreira *et al.* [Ferreira et al. 2009]. A mobilidade veicular e a comunicação entre os veículos são dois fatores fundamentais desta modelagem, que baseou-se no uso de fotografias aéreas (*stereoscopic aerial photography*) e mostrou-se mais realista que outros simuladores para VANETs [Ferreira et al. 2009]. De fato, a modelagem do tráfego urbano é uma tarefa árdua, sendo que cada cidade possui suas peculiaridades. Por exemplo, os motoristas preferem trafegar por determinadas vias em determinados horários, de acordo com o movimento dos pedestres e/ou de outros veículos nestes locais.

Nesta análise consideramos o grau de conectividade veicular desta cidade, onde as rotas escolhidas pelos veículos constituem o fator que mais interfere nesta métrica, pois a concentração de veículos em um número reduzido de vias certamente afetará a conectividade de veículos que trafegam em outras vias.

A Figura 5(a) apresenta o grau de conectividade entre dois veículos em relação à distância entre os mesmos (dados extraídos de [Ferreira et al. 2009]). Para este cálculo, fixou-se um veículo i e calculou-se a média de caminhos disjuntos nos nós que ligam os outros veículos ao i, de acordo com a distância que os mesmos se encontram. O gráfico apresenta o grau de conectividade para cenários com aproximadamente 80 e 240 veículos/km^2. Além disso, o alcance de transmissão foi especificado em $250m$ para comunicações diretas e $140m$ para comunicações onde alguma construção impediu a comunicação direta.

A partir destes dados, podemos derivar o número máximo de faltas toleradas na execução do BFT-CUP entre estes veículos: como é necessário $k \geq 2f + 1$ (Seção 2.4), onde k é o grau de conectividade, temos que $f \leq \frac{k-1}{2}$. A Figura 5(b) mostra que é possível tolerar até aproximadamente 40 faltas em um cenário de 240 veículos/km^2, que se encontram em uma distância de até $100m$. Além disso, este número diminui bruscamente até os $400m$ de distância quando então tende a estabilizar.



(a) Conectividade Veicular da Cidade de Porto (extraído de [Ferreira et al. 2009])

(b) Quantidade de Faltas Toleradas

Figura 5. O número de faltas toleradas está diretamente relacionado com o grau de conectividade do sistema.

Existem diversas aplicações para o consenso em redes VANETs. Por exemplo, existe uma proposta de substituição dos semáforos (infraestrutura localizada na via) pela execução de

um consenso entre veículos [Ferreira et al. 2010], que determinará qual destes veículos deve atravessar por primeiro em um cruzamento, implementando uma espécie de semáforo virtual. Esta abordagem visa otimizar o fluxo de veículos em cruzamentos, pois neste caso um veículo apenas irá parar no cruzamento quando realmente for necessário devido à presença de outros veículos competindo pelo cruzamento.

5. Trabalhos Relacionados

Existem alguns trabalhos na literatura que buscam analisar analiticamente o grau de conectividade apresentado por uma rede MANET em função de suas configurações [Bettstetter 2002, Santi 2005]. Estes trabalhos buscam soluções para controlar a topologia da rede e podem ser empregados na obtenção de uma rede com características (principalmente relacionadas com o grau de conectividade) que permitam a solução do BFT-CUP.

Além disso, encontramos alguns trabalhos que buscam analisar o comportamento do FT-CUP (consenso tolerante a faltas por *crash*) [Costa and Greve 2007, Costa et al. 2009] através de simulações em uma rede MANET. Este trabalho complementa os anteriores através da análise dos mais diversos cenários e considera protocolos que admitem a presença de participantes maliciosos, o que também exige maior conectividade no sistema [Alchieri et al. 2008].

6. Conclusões

Este trabalho apresentou resultados e análises envolvendo aspectos práticos do BFT-CUP. Primeiramente, os protocolos do BFT-CUP foram simulados para diversos cenários em uma rede MANET. A análise destes resultados torna possível avaliar a probabilidade do BFT-CUP admitir solução dada uma determinada configuração da rede. Além disso, uma análise acerca da realização do BFT-CUP em uma rede VANET foi apresentada, onde tomou-se como base o tráfego veicular da cidade de Porto-Portugal.

A partir dos dados discutidos neste artigo é possível identificar quais são os parâmetros e configurações do sistema que devem ser considerados quando estes protocolos forem utilizados no desenvolvimento de aplicações: *(1)* a escolha do parâmetro f é determinante para a convergência de BFT-CUP; *(2)* uma escolha adequada entre o alcance de transmissão e a densidade de nós no sistema também determinará a possibilidade de convergência de BFT-CUP.

Além disso, nas simulações realizadas conseguimos resolver o BFT-CUP com um detector de participação mais simples do que o k-OSR. Nas configurações onde o BFT-CUP admitiu solução, geralmente o sistema apresentou um grafo de conectividade por conhecimento G_{di} k-fortemente conexo (apenas uma componente k-fortemente conexa). Outro ponto a destacar é que a propriedade de acordo não foi violada em nenhum cenário analisado, i.e., sempre que os protocolos do BFT-CUP terminaram, os participantes acabaram decidindo pelo mesmo valor.

Agradecimentos

Gostaríamos de agradecer à professora Fabíola Greve pelas discussões sobre os protocolos do BFT-CUP. Além disso, Eduardo Alchieri, Luiz Tomelin e Joni Fraga são bolsistas do CNPq.

Referências

Alchieri, E. A. P., Bessani, A. N., da Silva Fraga, J., and Greve, F. (2008). Byzantine consensus with unknown participants. In *12th International Conference On Principles Of DIstributed Systems*, pages 22–40.

Aspnes, J. (2003). Randomized protocols for asynchronous consensus. *Distributed Computing*, 16(2-3):165–175.

Barr, R., Haas, Z. J., and van Renesse, R. (2004). Jist: Embedding simulation time into a virtual machine. In *Proceedings o EuroSim Congress on Modelling and Simulation*.

Barr, R., Haas, Z. J., and van Renesse, R. (2005). *Handbook on Theoretical and Algorithmic Aspects of Sensor, Ad hoc Wireless, and Peer-to-Peer Networks*, chapter Scalable Wireless Ad hoc Network Simulation, pages 297–311. CRC Press.

Ben-Or, M. (1983). Another advantage of free choice: Completely asynchronous agreement protocols (extended abstract). In *Proceedings of the 2nd Annual ACM Symposium on Principles of Distributed Computing*, pages 27–30.

Bettstetter, C. (2002). On the minimum node degree and connectivity of a wireless multihop network. In *Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing - MobiHoc'02*.

Bracha, G. (1984). An asynchronous $\lfloor (n-1)/3 \rfloor$-resilient consensus protocol. In *Proceedings of the 3rd ACM symposium on Principles of Distributed Computing*, pages 154–162.

Castro, M. and Liskov, B. (2002). Practical Byzantine fault-tolerance and proactive recovery. *ACM Transactions on Computer Systems*, 20(4):398–461.

Cavin, D., Sasson, Y., and Schiper, A. (2004). Consensus with unknown participants or fundamental self-organization. In *Proceedings of the 3rd International Conference on Ad hoc Networks and Wireless - ADHOC-NOW 2004*, pages 135–148.

Cavin, D., Sasson, Y., and Schiper, A. (2005). Reaching agreement with unknown participants in mobile self-organized networks in spite of process crashes. Technical Report IC/2005/026, EPFL - LSR.

Chandra, T. D. and Toueg, S. (1996). Unreliable failure detectors for reliable distributed systems. *Journal of the ACM*, 43(2):225–267.

Correia, M., Neves, N. F., and Veríssimo, P. (2006). From consensus to atomic broadcast: Time-free Byzantine-resistant protocols without signatures. *The Computer Journal*, 49(1).

Costa, V. F., Greve, F. G. P., , and Tixeuil, S. (2009). Consenso ft-cup em redes manets: Uma abordagem prática. In *Anais do 27º Simpósio Brasileiro de Redes de Computadores - SBRC 2009*.

Costa, V. F. and Greve, F. G. P. (2007). Aspectos práticos da realização do consenso ft-cup em redes móveis ad-hoc. In *Anais VIII Workshop de Teste e Tolerância a Falhas- WTF 2007*, Belém, PA, Brasil.

Dwork, C., Lynch, N. A., and Stockmeyer, L. (1988). Consensus in the presence of partial synchrony. *Journal of ACM*, 35(2):288–322.

Ferreira, M., Conceição, H., Fernandes, R., and Tonguz, O. K. (2009). Stereoscopic aerial photography: an alternative to model-based urban mobility approaches. In *Proceedings of the sixth ACM international workshop on VehiculAr InterNETworking - VANET'09*.

Ferreira, M., Fernandes, R., Conceição, H., Viriyasitavat, W., and Tonguz, O. K. (2010). Self-organized traffic control. In *Proceedings of the seventh ACM international workshop on VehiculAr InterNETworking - VANET'10*.

Fischer, M. J., Lynch, N. A., and Paterson, M. S. (1985). Impossibility of distributed consensus with one faulty process. *Journal of the ACM*, 32(2):374–382.

Greve, F. G. P. and Tixeuil, S. (2007). Knowledge connectivity vs. synchrony requirements for fault-tolerant agreement in unknown networks. In *Proceedings of the International Conference on Dependable Systems and Networks - DSN*, pages 82–91.

Kotzanikolaou, P., Mavropodi, R., and Douligeris, C. (2005). Secure multipath routing for mobile ad hoc networks. *Wireless On-demand Network Systems and Services - WONS*, pages 89–96.

Lamport, L., Shostak, R., and Pease, M. (1982). The Byzantine generals problem. *ACM Transactions on Programing Languages and Systems*, 4(3):382–401.

Santi, P. (2005). Topology control in wireless ad hoc and sensor networks. *ACM Comput. Surv.*, 37:164–194.