

DIRESC: Um protocolo para descoberta e recuperação de dados em redes centradas em conteúdo e tolerantes a atraso

Cláudio Diego Souza¹, Danielle L. Ferreira¹, Carlos A. V. Campos¹

¹Programa de Pós-Graduação em Informática (PPGI)
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)
Rio de Janeiro – RJ – Brasil

{diego.souza, danielle.ferreira,beto}@uniriotec.br

Abstract. *Content-centric networks (CCN) has emerged as an innovative paradigm that modifies the traditional host-centric and address-based communication model. Its application to DTN scenarios is known to positively impact some network metrics and characteristics, such as delay and intermittence. Thus, in this paper we justify the feasibility of this combination and present DIRESC, a protocol for selection of message relays for data announcement, discovery and retrieval through a scheme of social cooperation between nodes. We implement our proposal using the ONE simulator and compared its performance against CCN broadcasting and PIFP protocols. In our results, DIRESC presented a decrease of at least 73.3% in data recovery delay and a gain of at least 10.1% in satisfaction of content requests over PIFP. Thus, we present DIRESC as a promising solution for accelerating the process of obtaining data in this type of network.*

Resumo. *Redes centrada em conteúdo (CCN) emergiram como um paradigma inovador que modifica o modelo tradicional de comunicação centrada no host e endereçamento. Sua aplicação aos cenários de DTN é conhecida por ter impacto positivo em algumas métricas e características de rede, como atraso e intermitência. Assim, neste artigo¹ justificamos a viabilidade desta combinação e apresentamos o DIRESC, um protocolo para a seleção de retransmissores de mensagem para o anúncio, descoberta e recuperação de dados através de um esquema de cooperação social entre nós. Implementamos nossa proposta usando o simulador ONE e comparamos seu desempenho com os protocolos CCN broadcasting e PIFP. Em nossos resultados, DIRESC apresentou uma redução de pelo menos 73,3% no atraso na recuperação de dados e um ganho de pelo menos 10,1% na satisfação dos pedidos de conteúdo sobre o PIFP. Assim, apresentamos o DIRESC como uma solução promissora para acelerar o processo de obtenção de dados neste tipo de rede.*

1. Introdução

Redes tolerantes a atraso ou interrupções (*Delay/Disruption Tolerant Networking* - DTN) preveem a existência de cenários de rede sem fio sem infraestrutura e de múltiplos saltos nos quais não se garante a existência de caminhos fim-a-fim entre o par origem-destino. Trata-se de ambientes de rede auto-organizáveis, nos quais tais dispositivos se comunicam

¹Os dois primeiros autores deste trabalho recebem bolsa de estudos da CAPES.

de maneira oportunista, ou seja, valendo-se de oportunidades de contato entre si para encaminhar pacotes ao longo da rede [Conti et al. 2015]. Neste ambiente, todos os nós são dispositivos finais, como celulares, tablets, notebooks etc, no entanto, se comportam como roteadores armazenando os pacotes recebidos e aguardando a oportunidade de retransmití-los a outros nós.

Em razão das características destas redes, como a alta dinamicidade das topologias, intermitência e baixa conectividade [Conti et al. 2015], os protocolos de roteamento tradicionais da arquitetura TCP/IP não se aplicam neste contexto, isto é, não é funcional a implementação de protocolos que preveem comunicações fim-a-fim em cenários nos quais ela é inexistente. Portanto, soluções de rede alternativas para tais ambientes têm tentado fugir de abordagem centradas em endereçamento para abordagens centradas em dados [Amadeo et al. 2014].

Uma solução possível é a utilização da arquitetura de redes centradas em conteúdo (*Content-Centric Networking - CCN*), conceitualmente introduzida por [Jacobson et al. 2009], que tem emergido como uma promissora abordagem de rede para a recuperação de dados a partir da desconstrução do conceito de endereçamento no processo de busca por provedores de serviços. Diferentemente das redes centradas em hospedeiro (*host-centric*), CCN foca na identificação do conteúdo buscado, modificando o paradigma de comunicação para uma rede centrada em dados (*data-centric*).

Em CCN, a comunicação é conduzida pelos nós consumidores ou requisitantes que, quando interessados em determinado conteúdo, disseminam pacotes de interesse na rede. Quando estes pacotes atingem um nó que possua o conteúdo buscado, tal nó envia os pacotes de dados ao requisitante, concluindo o processo de busca e entrega. Além do encaminhamento de interesses baseado em nomes de conteúdos, o paradigma CCN se caracteriza também pela replicação de pacotes de dados em cache na rede, o que facilita a entrega destes pacotes ao requisitante [Jacobson et al. 2009].

Os nós que compõem um cenário CCN possuem três componentes básicos [Jacobson et al. 2009]: a *content store* (CS), que é a cache para armazenamento de dados; a *forwarding information base* (FIB), uma tabela que mapeia os anúncios de provedores de conteúdo à interface pela qual tal anúncio foi recebido; e a *pending interest table* (PIT), uma tabela que registra as requisições de conteúdos realizadas e ainda não satisfeitas, mapeando-as às interfaces por onde tais requisições chegaram.

Em [Liu et al. 2016] são destacados alguns benefícios que o paradigma CCN é capaz de prover ao paradigma DTN. Dentre eles, destaca-se o fato de CCN permitir comunicações assíncronas entre um par requisitante-provedor, isto é, sem a necessidade de que estes nós estabeleçam e mantenham uma comunicação fim-a-fim. Além deste, a busca de dados por nomes facilita a sua recuperação em ambientes de alta mobilidade, quando não se sabe a localização do nó, como ocorre em DTN. Por fim, a replicação de dados na rede reduz o número de requisições e o atraso na recuperação destes dados.

Apesar dos benefícios que a aplicação CCN é capaz de prover no contexto de DTN, a sua combinação levanta questões de pesquisa que ainda precisam ser estudadas. Em DTN, um desafio chave de roteamento é determinar a estratégia apropriada de seleção do nó retransmissor a fim de aumentar a probabilidade de entrega de pacotes, minimizando o número de replicações na rede e reduzindo ao máximo o atraso desta en-

trega. O protocolo Epidêmico [Becker et al. 2000], que é baseado no encaminhamento de pacotes em toda oportunidade de contato, apresenta bons resultados em termos de atraso e taxa de entrega, mas acarreta um alto custo em replicações de pacotes e, portanto, sobrecarga na rede [Zhu et al. 2015]. Já as comunicações em CCN foram idealizadas para operar em *broadcast* [Jacobson et al. 2009]. Uma tendência nos trabalhos mais recentes que utilizam DTN e CCN conjuntamente é encontrar um meio termo entre a política de funcionamento de um (neste caso, CCN, conceitualmente projetado para operar em *broadcast*), às limitações de outro (DTN, que precisa evitar sobrecarga e exaustão de recursos nos nós), e assim tornar os processos de encaminhamento mais “conscientes” por meio de estratégias como a seleção de retransmissores, a fim de evitar o comportamento epidêmico na comunicação entre os nós.

A estratégia de seleção de nós retransmissores para a descoberta e recuperação de conteúdo se propõe a limitar o número de nós que recebem pacotes por meio de uma escolha baseada em algum critério. Com base nesta estratégia, no presente artigo propomos o DIRESC, um protocolo para anúncio, descoberta e recuperação de dados em redes centradas em conteúdo e tolerantes a atraso (*Content-Centric Delay Tolerant Network - CCDTN*) baseado na cooperação entre nós que possuem determinado grau de relacionamento social.

Nossa proposta se baseia na seleção de nós retransmissores, uma vez que, por meio dela, evita-se o envio de mensagens a nós de modo indiscriminado na rede, o que apresenta um impacto significativo na quantidade de pacotes replicados e transmissões realizadas. Assim, recursos como energia e capacidade de *buffer* podem ser poupados. O esquema de cooperação social proposto se baseia no relacionamento existente entre os nós, calculado por meio do seu histórico de encontros. Este relacionamento é capaz de influenciar a probabilidade de entrega de mensagens, em vista que permite inferir se um par de nós específico possui chances de se reencontrar na rede, qualificando-os ou não como nós aptos à retransmissão, o que influencia também no atraso de recuperação de conteúdo. Nossa proposta define três possibilidades de envio de requisições e outras três para entrega de conteúdo, diferentemente de outras abordagens que, além de dar ênfase no processo de descoberta (envio de interesses), dando menor destaque ao processo de entrega de dados, definem um único critério para determinar a competência de um nó em entregar um pacote. Assim, utilizamos a estratégia de seleção de retransmissores de maneira mais distribuída e, portanto, sem exigir intensamente de apenas um conjunto de nós específico na rede.

Assim, nossa proposta se divide em três etapas principais para o funcionamento da rede: anúncio, descoberta e recuperação de conteúdo. Cada uma delas leva em consideração o fator de relacionamento social entre nós para o seu funcionamento e, de modo geral, visam reduzir a distância (temporal e em número de saltos) entre um requisitante e o conteúdo solicitado.

As principais contribuições deste trabalho podem ser resumidas da seguinte forma:

1. Proposta do protocolo DIRESC para anúncio, descoberta e recuperação de conteúdo em CCDTN. O protocolo proposto obteve um aumento no percentual de satisfação de requisições, ou seja, na relação entre interesses enviados e atendidos, mostrando resultados melhores quando comparado com outras propostas. Além disso, o DI-

RESC apresentou um alto decaimento no atraso médio da recuperação de dados quando comparamos nossa proposta com outras abordagens.

2. A implementação da arquitetura CCN e do protocolo DIRESC no ambiente de simulação ONE (*Opportunistic Network Environment*), o qual foi projetado para representar redes toletantes a atraso.

O restante do artigo está estruturado da seguinte forma. A Seção 2 aborda trabalhos anteriores relacionados a nossa proposta. A Seção 3 descreve o funcionamento do protocolo DIRESC. A Seção 4 apresenta a avaliação de desempenho realizada sobre o protocolo proposto e uma comparação com outros trabalhos da literatura. Por fim, a Seção 5 apresenta as conclusões e trabalhos futuros.

2. Trabalhos relacionados

No que concerne a combinação de CCN com redes sem fio sem infraestrutura e de múltiplos saltos, [Kim and Ko 2015] elabora um mecanismo proativo de atualização periódica de informações sobre nós da rede, e outro que propõe a transmissão de mensagens para recuperação confiável de conteúdo baseada na classificação destes nós, além de fazer uso da política de seleção de nós retransmissores, que neste caso se baseia na distância em número de saltos até o destino e qualidade do enlace. A proposta atinge 50% de ganho em taxa de transferência quando comparada com o método básico de encaminhamento oportunista por *broadcast*.

Por sua vez, em [Anastasiades et al. 2014] é apresentada uma proposta baseada em agente, na qual um nó requisitante, por meio de um processo de sondagem e comunicação *handshake* de três vias, encontra nós agentes em potencial (de acordo com informações como histórico de locais visitados) e delega a um deles a função de descoberta e recuperação de conteúdo em uma rede CCN oportunista. Os resultados do trabalho apontam que um intervalo de sondagem de 30s apresenta o melhor desempenho no cenário proposto em vista que o número de notificações pode ser diminuído em até 80%.

O trabalho de [Lu et al. 2014] propõe o STCR (*Social-Tie based Content Retrieval*) que, por meio do cálculo de relacionamento e centralidade social e do algoritmo de *clustering K-means*, agrupa em *clusters* nós de centralidade social semelhante. Neste contexto, anúncios de provedores de conteúdo e requisições convergem para os nós de maior centralidade social. Já o envio destas requisições (de um nó de mais alta centralidade até um provedor) e os de dados (do provedor ao requisitante) é realizado através de nós que possuam alto relacionamento social com o destinatário do pacote em questão. A proposta apresenta ganhos em termos de atraso, custo e taxa de entrega em relação aos protocolos *Epidemic*, *BubbleRap* e *PeopleRank*. Por outro lado, o STCR, por seccionar o processo de busca por dados em duas etapas (enviar intererres primeiramente para os nós de mais alta centralidade e só então enviá-los para os provedores de conteúdo), aumenta o tempo de recuperação de dados em comparação com outros mecanismos para CCN aplicada no cenário DTN, como é verificado em [Duarte et al. 2015].

Baseado no protocolo PROPHET, de roteamento em DTN, [Duarte et al. 2015] elabora o PIFP (*Probabilistic Interest Forwarding Protocol*), um protocolo probabilístico de encaminhamento de pacotes de interesse em DTN, a partir do qual os nós calculam a probabilidade de encontrarem determinado conteúdo e utilizam esta informação para selecionar retransmissores de interesses. Isto é, nós requisitantes de conteúdo apenas

enviam interesses para outros nós que possuam uma probabilidade superior a sua. Já o envio de pacotes de dados possui duas variantes: no PIFP, este envio considera apenas o caminho reverso ao envio dos interesses; já no PIFP *Proative*, os dados são enviados a qualquer nó da rede. O trabalho apresenta ganhos em termos de número de interesses enviados e atendidos e atraso de resposta em relação ao STCR. No entanto, por convocar sempre os nós com maiores probabilidades, o PIFP negligencia a competência de outros nós, oferecendo uma seleção não balanceada de retransmissores.

Em uma análise mais acurada dos trabalhos supracitados, podemos levantar uma questão: suas abordagens não estariam limitando tão fortemente as replicações de interesses a ponto de prejudicar a rede em termos de atraso de entrega e percentual de satisfação de requisições? Além disso, dado o ranqueamento dos “melhores” nós da rede segundo algum critério, seja centralidade social ou probabilidade de encontro com conteúdo, a constante seleção dos nós mais bem posicionados no ranqueamento produz uma “distribuição injusta de carga”, ou seja, faz com que a maior parte do tráfego da rede seja realizada apenas por uma pequena parcela de nós, sobrecarregando-os e exaurindo rapidamente seus recursos, como explica [Junior et al. 2014].

3. O protocolo DIRESC

Visando superar os desafios da combinação de DTN e CCN e a questão levantada na seção anterior, propõe-se o DIRESC, um protocolo para descoberta e recuperação de conteúdo baseado em cooperação social, que tem por objetivo principal reduzir o atraso de descoberta e recuperação de conteúdo em redes CCDTN. Diferentemente de outras abordagens propostas na literatura, nas quais os conteúdos são alcançados através da convergência de pacotes de interesse por meio do que [Junior et al. 2014] chama de “escolha de nós preferidos”, uma vez que definem um único critério para determinar a competência de um nó em entregar um pacote, o DIRESC propõe uma abordagem distribuída, evitando assim tal comportamento indesejável, resultante do funcionamento de protocolos oportunistas com base social. A definição de três possibilidades de envio de requisições e outras três para entrega de dados evita o comportamento em *broadcast*, típico de CCN, mas torna a proposta mais tolerante em termos de critérios de seleção de nós retransmissores.

3.1. Relacionamento social

O fator de relacionamento social $R_i(j)$, entre o par de nós i e j , é uma medida calculada de acordo com o histórico de encontros entre os mesmos. Para a definição do seu valor, leva-se em consideração o quão recente os encontros ocorreram e a sua quantidade, ou seja, a frequência com a qual estes nós entram em contato. O valor de $R_i(j)$, baseado em [Lu et al. 2014], é definido da seguinte forma.

$$R_i(j) = \sum_{k=1}^N F(t_{base} - t_{jk}) \quad (1)$$

A influência de cada encontro entre os nós i e j , sendo N o total de encontros entre eles até o instante t_{base} , é determinado por $F(x)$, uma função que calcula o impacto gerado em $R_i(j)$ pelo quão recente estes encontros ocorreram, em que x é o intervalo de tempo entre um encontro no passado e o tempo atual. A função $F(x)$ é definida da seguinte maneira:

$$F(x) = \left(\frac{1}{2}\right)^{\lambda x}, \text{ onde } \lambda = e^{-4}. \quad (2)$$

O protocolo DIRESC utiliza o cálculo do relacionamento social $R_i(j)$ como princípio. No entanto, diferentemente de [Lu et al. 2014], que utiliza essa medida para a formação de agrupamentos entre nós que possuem valor de centralidade social similar, o DIRESC utiliza o valor de $R_i(j)$ como um parâmetro para determinar se um par de nós possui um alto relacionamento social entre si. Utilizamos uma abordagem diferenciada do relacionamento social uma vez que, por se valer da métrica de centralidade, [Lu et al. 2014], além de ocasionar o problema da escolha de “nós preferidos”, tem seu processo de descoberta de conteúdo seccionado, elevando o atraso de recuperação dos dados.

Portanto, o DIRESC compara $R_i(j)$ com a sua média, $R_{medio} = \frac{1}{N} \sum_k R_i(k)$, no qual $k = 1, 2, \dots, N$ e N é o número total de nós com os quais i teve contato até o instante da conexão com j . Logo, se $R_i(j) \geq R_{medio}$ então i e j possuem alto relacionamento social.

A fim de manter o fator de relacionamento social sempre atualizado, a todo contato entre dois nós, os mesmos atualizam o valor de $R_i(j)$ e, identificando que esses nós possuam um alto relacionamento social eles passam a ser considerados “amigos” e a tomar decisões que definirão o comportamento da rede. Para a melhor compreensão destas decisões, elas foram enquadradas em três fases, definidas como: (i) distribuição, (ii) busca e (iii) retorno.

3.2. Fase de distribuição

A fase de distribuição compreende a disseminação de informações acerca de conteúdos. Nela ocorrem dois tipos de anúncios: no primeiro, nós produtores, ou seja, nós que geram pacotes de dados, informam para nós com quem têm alto relacionamento social quais são os dados que possuem em CS; no segundo, nós que receberam anúncios de produtores comunicam a seus amigos tal recepção. Assim, delimitam-se agrupamentos capazes de responder às solicitações para tais conteúdos. Portanto, nesta etapa, assumindo que dois nós sejam amigos, segue-se as seguintes trocas de anúncios:

1. Anúncio direto: caso o nó i seja um produtor e possua conteúdos na CS não registrados na FIB de j , ou não esteja mapeado como um nó de próximo salto caso tais registros de fato existam, o nó j atualiza sua FIB, criando registros que façam o mapeamento entre o nó i e os prefixos que ele atende, ou inserindo tal nó nos registros existentes como um próximo salto para envio de interesses. Ou seja, o nó j passa a compreender que i é capaz de atender solicitações para determinada porção de conteúdos.

2. Anúncio transitivo: caso a FIB do nó i possua registros que não existem na FIB de j , esta é atualizada com os registros faltantes de modo que i seja mapeado como um nó de próximo salto para o envio de interesses.

Deste modo, uma vez que existe alto relacionamento social entre nós, criam-se agrupamentos capazes de responder às solicitações de forma ágil, tendo em vista do elevado nível de contatos entre os nós que os compõem. O intuito de um agrupamento é, portanto, reduzir o tempo de descoberta de conteúdos a partir do momento em que os pacotes de interesse o atingem. A Figura 1 sumariza as tarefas de anúncios no DIRESC.

3.3. Fase de busca

A fase de busca, por sua vez, diz respeito a requisição de conteúdo. Nela, selecionamos nós que se enquadrem em três critérios para envio de interesses, a fim de evitar que estes



Figura 1. Anúncio direto e transitivo

sejam transmitidos de forma epidêmica, mas dando certa tolerância para que tais envios ocorram. Leva-se em consideração informações que os nós precisam registrar acerca de requisições. Primeiramente, define-se dois contadores de retransmissões para cada requisição e um valor máximo que cada um pode atingir. Cada contador se designa a um grupo específico para o qual se pode enviar pacotes de interesse. São eles: o agrupamento de nós que possui informações em FIB sobre certo conteúdo; e nós de alto relacionamento social com o detentor do pacote de interesse. Assim, um nó, de posse de um interesse, somente pode enviá-lo até que o número de envios por grupo atinja o número máximo permitido, evitando que se eleve o número de replicações deste pacote exacerbadamente. Desta forma, os nós transmitem pacotes de interesse caso quaisquer das premissas a seguir seja verdadeira, como observado na Figura 2:

1. O nó encontrado possui os dados armazenados na CS. Neste caso, não se contabiliza o envio de interesse (ver Figura 2(a)).

2. O nó encontrado pertence ao agrupamento de nós que possui informações na FIB a respeito de tal conteúdo e o número de envios da requisição em questão para este grupo é menor que o valor máximo permitido. Neste caso, registra-se na PIT uma nova retransmissão da requisição pelo critério de agrupamento (ver Figura 2(b)).

3. Os nós possuem um alto relacionamento social entre si e o número de envios da requisição entre eles é menor que o valor máximo permitido. Neste caso, registra-se na PIT uma nova retransmissão da requisição pelo critério de amizade. Este critério de envio aumenta a persistência de pacotes de interesse em descobrir conteúdo na rede, de modo que os nós que se encaixam nesta avaliação trabalham de modo socialmente cooperativo buscando dados em favor do solicitante, tendo em vista que este par de nós possui grandes chances de se reencontrar (ver Figura 2(c)).

Além disso, dados os critérios supracitados para envio de pacotes de interesse, utiliza-se um esquema de priorização de envio, no qual um nó interessado em determinado conteúdo dá prioridade de envio de solicitações a nós que se enquadram nos critérios acima, na ordem em que eles se apresentam.

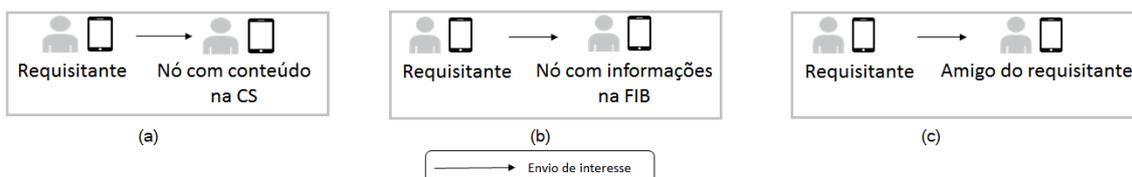


Figura 2. Possibilidades de envio de pacotes de interesse

3.4. Fase de retorno

Assim como na fase de busca, na fase de retorno, selecionamos nós que se enquadram em três critérios para envio de dados, evitando que estes sejam transmitidos de forma

epidêmica, e dando tolerância para o envio destes pacotes, aumentando suas chances de entrega. Esta fase se refere a entrega de conteúdo, ou seja, quando um nó que possui na CS o pacote de dados correspondente a uma solicitação recebida, ele enviará tais dados a outros nós baseado nos seguintes critérios, como observado na Figura 3:

1. Para todo nó que lhe fez tal solicitação, ou seja, todo nó que lhe enviou um pacote de interesse para os dados que ele possui (ver Figura 3(a)).

2. Para todo nó que possua na PIT um registro para aqueles dados e, portanto, é um solicitante, ainda que não tenha lhe enviado um pacote de interesse, mas está aguardando receber os dados por meio de outro nó (ver Figura 3(b)).

3. Para todo nó que possua um alto relacionamento social com o nó requisitante original dos dados (o nó que inseriu na rede o pacote de interesse recebido por este que possui os dados). Deste modo, mais uma vez, utiliza-se o valor obtido do cálculo de relacionamento social, no entanto, para auxiliar a recuperação dos dados (ver Figura 3(c)).

Da mesma forma que na fase de busca, utiliza-se um esquema de priorização de envio de acordo com a ordenação apresentada dos critérios para entrega de dados.

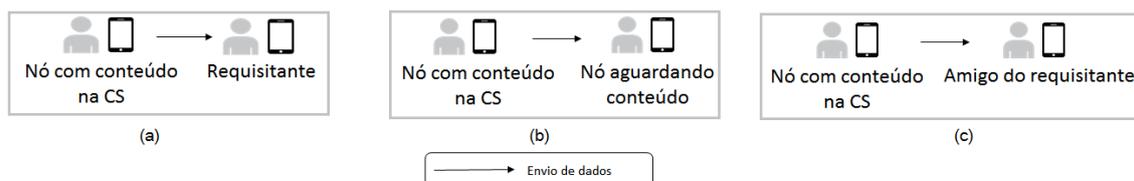


Figura 3. Possibilidades de envio de pacotes de dados

3.5. O algoritmo

O Algoritmo 1 descreve o funcionamento do DIRESC, resumindo as tomadas de decisão dos processos de anúncio, descoberta e recuperação de dados. Cada um destes processos, identificados nas iterações do algoritmo, respectivamente, é representado por uma das fases descritas anteriormente (distribuição, busca e retorno).

4. Avaliação de desempenho

Nesta seção serão apresentados os resultados obtidos para o protocolo DIRESC, que foi experimentalmente avaliado por meio do simulador ONE (*Opportunistic Network Environment*) [Keränen et al. 2009], no qual foi implementada a arquitetura de redes CCN, de modo que se tornasse possível combinar este modelo de comunicação com a proposta do simulador, que é voltada para redes tolerantes a atraso. Para a geração dos gráficos foi utilizado o *software* R [R Core Team 2016].

O objetivo de avaliar o desempenho do DIRESC é para verificar a sua capacidade de alcançar resultados satisfatórios em termos de taxa de recuperação de conteúdo e controle da sobrecarga de pacotes que são inseridos na rede, de modo que o atraso para a obtenção dos dados seja reduzido em comparação com as outras abordagens. Em seguida, o cenário de simulação será apresentado e os resultados experimentais serão mostrados e discutidos.

Algoritmo 1: DIRESC: PSEUDOCÓDIGO

```
Entrada:  $i, j, R_{medio}, \text{maxEnvios}$ 
1 início
2   para todo contato de  $i$  e  $j$  faça
3     calcula-se  $R_i(j)$  (ver Equação (1))
4     % Fase de distribuição do DIRESC
5     se  $R_i(j) \geq R_{medio}$  então
6       se  $i$  é produtor então
7         para cada prefixo  $p$  na CS de  $i$  que não existe na FIB de  $j$  faça
8           Atualize a FIB de  $j$ : adicionando o prefixo  $p$  e mapeando-o para  $i$ 
9         fim
10      fim
11      para cada registro  $r$  na FIB de  $i$  que não existe na FIB de  $j$  faça
12        Atualize a FIB de  $j$ : adicionar uma cópia deste registro  $r$  mapeando-o para  $i$ 
13      fim
14    fim
15    % Fase de busca do DIRESC
16    para cada entrada na PIT de  $i$  interessada no conteúdo  $c$  faça
17      se  $c$  está presente na CS de  $j$  então
18         $i$  envia um pacote de interesse em  $c$  para  $j$ 
19      fim
20      senão
21        se ( $\text{há um registro na fib de } j \text{ para } c \vee j \text{ é um nó de próximo salto para } c \text{ na FIB de } i$ )  $\wedge$  o
22          número de submissões desta requisição para o cluster  $< \text{maxEnvios}$  então
23             $i$  envia um pacote de interesse em  $c$  para  $j$ 
24            O número de submissões desta requisição para o cluster é aumentado
25          fim
26          senão
27            se  $R_i(j) \geq R_{avg} \wedge$  o número de envios desta requisição para amigos  $< \text{maxEnvios}$ 
28              então
29                 $i$  envia um pacote de interesse em  $c$  para  $j$ 
30                O número de submissões desta requisição para amigos é aumentado
31              fim
32            fim
33          fim
34        fim
35      % Fase de retorno do DIRESC
36      para cada interesse no conteúdo  $c$  recebido faça
37        se  $c$  está presente na CS de  $i$  então
38          se  $j$  é o requisitante  $\vee$  há uma entrada na PIT de  $j$  interessada em  $c \vee R_j(\text{requisitante}) \geq$ 
39             $R_{avg} \circ f_j$  então
40               $i$  envia o pacote de dados  $c$  para  $j$ 
41            fim
42          fim
43        fim
44      fim
45    fim
46  fim
```

4.1. Cenário de simulação

O cenário utilizado foi o mesmo descrito em [Duarte et al. 2015] e foi adotado por ser bastante diverso, com diferentes perfis de mobilidade, e para que pudéssemos comparar o DIRESC com o PIPF e CCN *broadcasting* de forma justa.

O cenário contempla a existência de 100 nós, dos quais 62 são pedestres, 32 são carros e 6 são bondes, se movimentando no mapa da cidade de *Helsinki*, Finlândia. Os nós se movem de acordo com diferentes modelos de mobilidade determinado para cada grupo de nós. Pedestres e carros se movem segundo o modelo *Shortest Path Map-Based Movement*, que utiliza o algoritmo Dijkstra para encontrar caminhos mais curtos entre dois pontos no mapa; enquanto que bondes usam o modelo *Map-Based Movement*, o qual é parametrizado por rotas específicas para determinar sua movimentação. A tecnologia utilizada foi o Bluetooth, com taxa de transmissão de 250kbps e alcance de 10m. A Tabela

1 apresenta mais detalhes acerca do cenário de simulação.

A criação das mensagens de interesse são agendadas durante a simulação por meio da classe de leitura de eventos *StandardEventsReader*, disponível no simulador ONE. Assim, através da leitura do *trace* que registra esses agendamentos, os pacotes de interesse são gerados no *buffer* de nós escolhidos aleatoriamente, em diferentes instantes de tempo, durante a simulação. A criação dos pacotes de interesse está sujeita a capacidade do *buffer* dos nós. Além disso, os nós enviam os pacotes de interesse gerados de acordo com os critérios estabelecidos pelos protocolos propostos. O DIRESC utiliza o procedimento descrito na Seção 3.3. O protocolo PIFP calcula a probabilidade que os nós possuem de encontrar determinado conteúdo e o envio de interesses é realizado para os aqueles que possuem o maior valor neste cálculo. Por sua vez, o protocolo CCN *broadcasting* se refere ao comportamento padrão da arquitetura CCN, isto é, por meio do envio de pacotes em toda oportunidade de contato.

De forma análoga, a geração dos pacotes de dados segue um procedimento similar ao descrito para as mensagens de interesse. Ou seja, os nomes de conteúdos são lidos a partir de um *trace* e os pacotes de dados correspondentes são criados na CS dos produtores. No total, são criados 500 conteúdos diferentes em 15 nós produtores igualmente definido em [Duarte et al. 2015].

Tabela 1. Parâmetros da simulação

Parâmetro	Valor
Duração	86400s (24h)
Área (metros x metros)	4500m x 3500m
Velocidade média (m/s)	Pedestres: 0.5 a 1.5, Carros: 2.7 a 13.9, Bondes: 10 a 30
Tempo de vida das requisições	8 horas
Capacidade de <i>buffer</i>	Consumidores: 50MB, Produtores: 100MB

Os resultados reportados refletem 30 rodadas de simulação para cada um dos protocolos (DIRESC, PIFP e CCN *broadcasting*). Para o cálculo das métricas apresentadas, considerou-se um intervalo de 15 minutos para contabilizar informações sobre trocas de mensagens, no qual cada observação assumiu-se o valor acumulado destas informações desde o intervalo anterior até o atual. O mesmo cenário foi utilizado para as simulações dos três protocolos supracitados.

4.2. Resultados experimentais

O protocolo DIRESC foi avaliado em termos do atraso médio e da taxa de recuperação de conteúdo, bem como do controle de sobrecarga de pacotes que trafegam pela rede. Os resultados mostram que o protocolo proposto é capaz de reduzir significativamente, cerca de 70%, o atraso médio necessário para a obtenção do conteúdo em relação a outras abordagens, com um menor tráfego de mensagens.

Com o objetivo de avaliar o protocolo DIRESC e apresentar os resultados, foram definidas as seguintes métricas:

Relação entre interesses enviados e satisfeitos: refere-se à relação absoluta e percentual entre a quantidade de conteúdo entregue aos requisitantes e o número de requisições feitas por estes. Deste modo, pode-se analisar o quão satisfatória uma proposta é em relação à quantidade de solicitações de conteúdos feitas na rede.

Carga total: definida como quantidade de pacotes de interesse e dados inseridos na rede, na qual se observa a quantidade de replicações de pacotes feita na rede.

Atraso de recuperação de conteúdo: refere-se ao intervalo de tempo médio calculado para todas as recuperações de dados, desde o momento do envio da solicitação até a entrega de conteúdo. Assim, avalia-se o quanto uma proposta é capaz de reduzir o tempo de descoberta e recuperação de dados na rede.

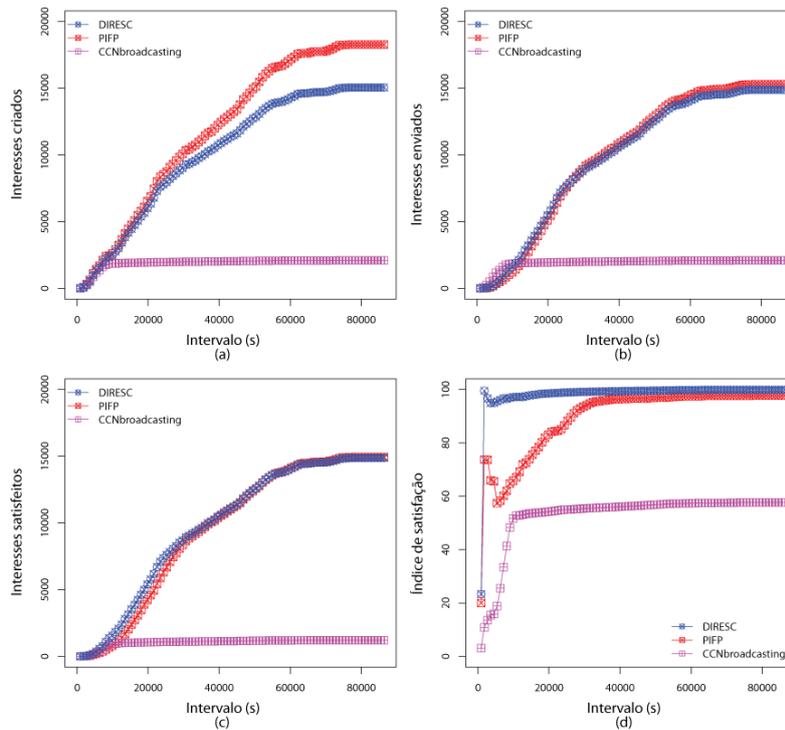


Figura 4. Relação entre interesses enviados e satisfeitos

A Figura 4 apresenta a relação entre os interesses criados, enviados e satisfeitos na rede. Considera-se uma mensagem de interesse satisfeita quando o nó requisitante que gerou uma mensagem de interesse recebe o conteúdo requisitado. A quantidade de mensagens de interesse criadas na rede é mostrada na Figura 4(a). Nela, observa-se que esta quantidade é semelhante para os protocolos DIRESC e PIFP, diferentemente do protocolo CCN *broadcasting* que exhibe uma quantidade bem reduzida das mensagens de interesse criadas. Esse comportamento pode ser explicado pela característica do protocolo CCN *broadcasting*, na qual a disseminação das mensagens de interesse é feita a todo contato entre nós. Dessa forma, o protocolo inunda a rede com mensagens de interesse, exaurindo a capacidade do *buffer* rapidamente e levando ao comportamento mostrado.

As transmissões dos pacotes de interesse estão sujeitas aos critérios de cada protocolo, portanto, observa-se na Figura 4(b) que as curvas de envio são bem semelhantes as de criação de interesses. O objetivo deste gráfico é complementar a observação da Figura 4(c), de modo que se possa comparar a quantidade de requisições efetuadas e quantas delas foram atendidas.

A Figura 4(c) apresenta a quantidade de requisições originais, ou seja, feitas por nós que deram origem a pacotes de interesse, que foram satisfeitas. Uma vez que as curvas se apresentam de maneira muito semelhante as dos gráficos anteriormente tratados, complementamos a análise desta métrica com a Figura 4(d), que apresenta a relação

percentual entre interesses enviados e satisfeitos. O DIRESC apresenta um melhor desempenho em termos de satisfação em relação aos demais. Ainda que o PIFP se aproxime de nossa proposta a partir de certo ponto, ainda obtivemos um melhor resultado até 30000 segundos. Além disso, nosso esquema de priorização de envio garante que, no início dos experimentos, os interesses sejam enviados para nós com conteúdo na CS, uma vez que as informações de relacionamento social e de anúncios ainda estão sendo preparadas.

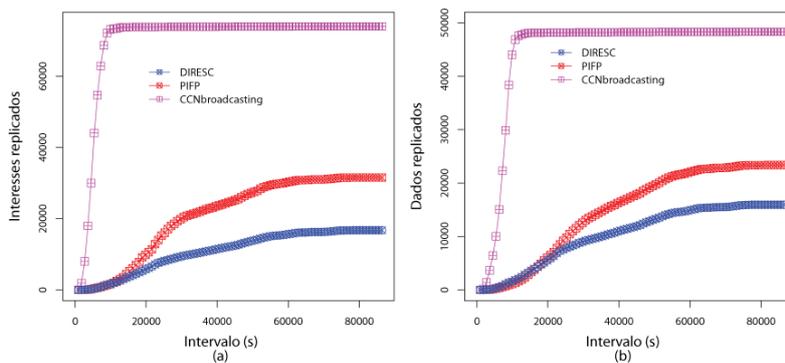


Figura 5. Carga total de pacotes de dados e interesses replicados na rede

As Figuras 5(a) e 5(b), por sua vez, apresentam o quanto de carga de mensagens é gerado na rede. Contabilizamos aqui a quantidade de réplicas de pacotes de dados e de interesse que foram feitas por retransmissões. Observa-se por meio deste gráfico que o DIRESC onera muito menos a rede em termos de sobrecarga em relação aos demais protocolos. Isto acontece devido ao esquema de priorização de envio, que, no caso dos interesses, antepõe nós que possuam o conteúdo ou qualquer informação sobre ele e, em último caso, privilegia nós “amigos”, mesmo que isso também auxilie a recuperação de conteúdo e em reduzir o seu atraso; já no caso dos dados, a priorização privilegia qualquer nó interessado em obter o conteúdo, deixando para última instância nós que apenas possam fazer esta entrega. Assim, dispendo nós candidatos a serem retransmissores no último grau do critério de priorização, reduzimos o número total de réplicas na rede.

Nossa última métrica a ser analisada, o atraso de recuperação de conteúdo, compreende a média dos intervalos de tempo cumulativos obtidos desde o momento do envio de um pacote de interesse até o instante no qual o nó requisitante original recebe o pacote de dados que o satisfaça. Em vista das muitas replicações de mensagens e, portanto, do elevado preenchimento do *buffer*, os nós que operam por meio de *CCN broadcasting* estão mais propensos a recusar o recebimento de mensagens por falta de espaço neste recurso.

Devido a isto, podemos observar na Figura 6, um maior atraso (*CCN broadcasting*) até que os nós solicitantes possam receber os conteúdos requisitados, já que tanto interesses quanto dados podem ser largamente rejeitados no caminho entre origem e destino. Por sua vez, além do problema identificado de convergência de pacotes até o destino por meio de um único critério de envio, por considerar apenas o caminho reverso de envio dos pacotes de interesses para entregar os pacotes de dados, o PIFP ignora o fato de que contatos possam não se repetir, o que em DTN é muito provável, e, portanto, o conteúdo necessita ser entregue por outros caminhos. À vista disso, aumentam-se as chances de um nó ter de reexpressar um determinado interesse, o que justifica o elevado número de réplicas de interesses da Figura 5(b). Assim, compreende-se que o PIFP apresenta um

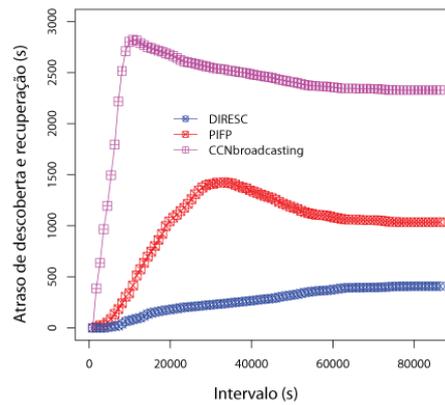


Figura 6. Atraso de recuperação de conteúdo

elevado atraso de recuperação de conteúdo por não propor um mecanismo de entrega diferenciado (ainda que o PIFP *Proactive* se proponha a isto, mas caia no erro de inundação da rede pelo envio exacerbado de pacotes de dados). O DIRESC, por outro lado, por propor três possibilidades de envio de interesses para descoberta e três possibilidades de envio de dados para recuperação de conteúdo, e definir um esquema de ordenação por prioridades entre eles, aumenta significativamente as chances de um nó requisitante recuperar um dado. Neste contexto, não apenas os nós detentores de conteúdo ou de informação sobre ele, mas os nós de alto relacionamento social com os requisitantes são capazes de operar de modo socialmente cooperativo, em favor destes requisitantes tanto no processo de descoberta quanto na recuperação daquele conteúdo. Como observado na Figura 6, este esquema foi capaz de reduzir consideravelmente o tempo de atraso de recuperação de conteúdo em relação aos outros protocolos.

Para finalizar esta seção, a Tabela 2 sumariza de forma percentual os ganhos obtidos, e observados nos gráficos, do protocolo DIRESC primeiramente em relação à CCN *broadcasting* apresentando ganhos expressivos de no mínimo 45% e, em seguida, em comparação ao protocolo PIFP apresentando melhor desempenho, principalmente, nas métricas de carga e atraso.

Tabela 2. Ganhos percentuais observados

Protocolos	Taxa de satisfação(%)	Carga (nº dados)	Carga (nº interesses)	Atraso (s)
DIRESC x CCN <i>broadcasting</i>	45,3%	76,9%	85,29%	89,07%
DIRESC x PIFP	10,07%	30,43%	50,45%	73,33%

5. Conclusões

Neste artigo, propusemos o DIRESC, um protocolo para anúncio, descoberta e recuperação de dados em CCDTN baseado em cooperação social. Nossa proposta utiliza o cálculo de relacionamento social como fator principal para tomada de decisões. Ao apresentar esquemas diferenciados de envios de interesses e dados, nossa abordagem apresentou resultados significativamente melhores em comparação ao comportamento CCN padrão, isto é, em *broadcasting*, e ao protocolo PIFP, que é baseado em probabilidade de encontros entre nós e conteúdos.

Os resultados foram obtidos por meio de experimentos realizados no simulador ONE, no qual a proposta foi implementada, juntamente com a arquitetura CCN. As

medições apresentadas neste trabalho mostram que o DIRESC é capaz de acelerar o processo de obtenção de dados em redes CCDTN e alcançar uma elevada taxa de satisfação de requisições em comparação as outras propostas.

Como trabalho futuro, planejamos aprimorar o funcionamento do DIRESC para que realize gerenciamento de cache, isto é, seja capaz de lidar com políticas de substituição de dados na CS a fim de gerir e utilizar da melhor forma este recurso.

Referências

- Amadeo, M., Campolo, C., Molinaro, A., and Ruggeri, G. (2014). Content-centric wireless networking: A survey. *Computer Networks*, 72:1–13.
- Anastasiades, C., El Alami, W. E. M., and Braun, T. (2014). Agent-based content retrieval for opportunistic content-centric networks. In *WWIC*, pages 175–188.
- Becker, V. D. et al. (2000). Epidemic routing for partially connected ad hoc networks. *Proceedings of Technique Report, Department of Computer Science, Duke University, Durham, UK*.
- Conti, M., Boldrini, C., Kanhere, S. S., Mingozzi, E., Pagani, E., Ruiz, P. M., and Younis, M. (2015). From MANET to people-centric networking: Milestones and open research challenges. *Computer Communications*, 71:1–21.
- Duarte, P., Macedo, J., Costa, A. D., Nicolau, M. J., and Santos, A. (2015). A Probabilistic Interest Forwarding Protocol for Named Data Delay Tolerant Networks. *7th AdhocNets 2015 San Remo, Italy, September 1-2, 2015 proceedings*, 155:94–107.
- Jacobson, V., Smetters, D. K., Thornton, J. D., Plass, M. F., Briggs, N. H., and Braynard, R. L. (2009). Networking named content. In *5th CoNEXT*, pages 1–12. ACM.
- Junior, N. M., Campos, C. A. V., and Lucena, S. C. (2014). Uma proposta de solução para o problema de nós preferidos em protocolos oportunistas com base social. In *Anais do 32º SBRC*, pages 91–104.
- Keränen, A., Ott, J., and Kärkkäinen, T. (2009). The ONE Simulator for DTN Protocol Evaluation. In *SIMUTools '09: Proceedings of the 2nd International Conference on Simulation Tools and Techniques*, New York, NY, USA. ICST.
- Kim, D. and Ko, Y.-B. (2015). A novel message broadcasting strategy for reliable content retrieval in multi-hop wireless content centric networks. *9th IMCOM '15*, pages 1–8.
- Liu, X., Li, Z., Yang, P., and Dong, Y. (2016). Information-centric mobile ad hoc networks and content routing: A survey. *Ad Hoc Networks*, 0:1–14.
- Lu, Y., Li, X., Yu, Y.-T., and Gerla, M. (2014). Information-centric delay-tolerant mobile ad-hoc networks. In *INFOCOM WKSHPs, IEEE Conference*, pages 428–433. IEEE.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Zhu, K., Li, W., Fu, X., and Zhang, L. (2015). Data routing strategies in opportunistic mobile social networks: Taxonomy and open challenges. *Computer Networks*, 93:183–198.