

Prediction-based redundant data elimination with content overheard in wireless networks

Apresentação I

Rafael Gonçalves de Oliveira Viana¹

13 de outubro de 2017

¹Sistemas de Informação – Universidade Federal do Mato Grosso do Sul (UFMS)
Caixa Postal 79400-000 – Coxim – MS – Brazil

rafael.viana@aluno.ufms.br

Resumo. Este artigo pretende melhorar a taxa de transferência de rede sem fio por supressão de transmissões de dados duplicados de links de rede. Demonstrou-se que a redundância da camada IP e a eliminação (RE) com o conteúdo excessivo pode significar melhorar o bom rendimento e a utilização de canais sem fio em dispositivos sem fio meio ambiente. No entanto, a integração da camada IP RE e a transmissão sem fio introduz um desafio. Ou seja, probabilístico ouvir por acaso sem fio e a possibilidade de um receptor ouvir de transmissores múltiplos causam caches de um remetente e um receptor longe da sincronização, que pode perturbar a camada de IP. A correção de RE e degrada seu desempenho. Os trabalhos anteriores lidam com este desafio pela probabilidade de ouvir por acaso que, no entanto, não é eficiente ou escalável. Nesse papel, propomos uma eliminação de redundância baseada em previsão com método de ouvir por acaso de conteúdo (PRECO) para resolver este desafio. Ao explorar o RE baseado em previsão, PRECO não requer sincronização de cache e estimativa de probabilidade de ouvir por acaso, o que permite a sua implantação eficiente e escalável. Baseado em PRECO, exploramos os benefícios da implantação do nível de sub-pacote RE como um serviço primitivo de camada IP em todos os nós na malha sem fio redes, propondo um protocolo de roteamento com reconhecimento de redundância. A avaliação de desempenho orientada por rastreamento mostra a eficácia e eficiência de PRECO em comparação com outros métodos RE.

1. Introdução

As redes sem fio têm sido uma comunicação amplamente utilizada paradigma para proporcionar mobilidade, Internet na cidade conectividade e computação ao ar livre com baixo custo e rápida implantação [1], [2], [3], [4], [5], [6]. No entanto, interferência e a qualidade da ligação fraca limita severamente o rendimento da rede sem fio redes especialmente para grandes redes densas [7], [8]. Este O papel concentra-se em uma importante classe de técnicas [9], [10], [11], [12] que visam melhorar o rendimento da rede sem fio suprimindo transmissões de dados duplicados da rede links. Essas técnicas podem eliminar as transmissões de pacotes ou conteúdos de dados que foram previamente transmitidos. Conseqüentemente, eles podem ser classificados em duas categorias: pacotes baseados Eliminação de redundância (RE) [9] e baseada em conteúdo RE [10], [11], [12],

[13]. Comparado com RE com pacotes, RE baseada em conteúdo pode explorar a localidade de dados na carga de trabalho transferido através de redes sem fio. A localidade dos dados é resultado do usuário acessar o mesmo conteúdo popular no Internet e também a similaridade entre diferentes Objetos de dados da Internet [10].

Estudos recentes mostraram que a grande maioria do tráfego A redundância na Internet decorre de trocas de dados duplicados de tamanho inferior a 150 bytes [14]. Por sua vez, o RE de camada IP em um nível de sub-pacote foi mostrado para fornecer informações significativas benefícios de desempenho em redes com fio. Faz uso de caches sincronizados entre o remetente e o receptor [15], [16], [17], [18]. O remetente remove a sequência de bytes duplicada (tão pequeno quanto 64B) da carga útil do pacote, comparando-o contra pacotes previamente transmitidos e, em vez disso, insere uma barra; o receptor substitui o calço pela sequência de bytes correspondente em pacotes recebidos anteriores para reconstruir o pacote completo.

Explorando a técnica RE da camada IP com conteúdo sobre-ouvir em redes sem fio pode melhorar significativamente a Goodput e utilização de canais sem fio [12]. Contudo, Isso apresenta um desafio. Alar-me sem probabilidade sem fio e a possibilidade de um receptor ouvir de múltiplas os transmissores causam os caches de um remetente e um receptor distante da sincronização, o que pode interromper a correção da RE e degradar sua performance. REfactor [12] trata da desafio pela estimativa de probabilidade de overhearing. Considerando um infra-estrutura de ponto de acesso único (AP) com múltiplos associados clientes, o remetente AP estima se o cliente receptor é provável que tenha ouvido um pedaço de saída do anterior transmissões para outros clientes. Com base nisso, o remetente calcula a redução esperada do tempo de transmissão resultante da remoção do pedaço, em consideração de possíveis Tempo adicional para solicitação de dados ausentes e retransmissão em caso de falta de cache no receptor. AP remove apenas o pedaço quando a redução esperada do tempo de transmissão é alta.

A probabilidade de overhearing sem fio, no entanto, é difícil estimar com precisão. REfactor determina o problema de overhearing capacidade apenas pela taxa de transmissão de dados, que é simples mas muito provavelmente não seja suficientemente preciso por causa da dinâmica condições de canal e interferência em grandes redes sem fio redes. Uma estimativa insuficientemente precisa pode causar erros decisões sobre remoção ou não de redundância, e conseqüências Reduza o desempenho da RE. Além disso, para estender para várias infra-estruturas AP e redes de malha, REfac- O tor exige que os nós sem fio enviem uma auditoria periodicamente notificações e comunicar conteúdo do cache, que incorre custo significativo de comunicação e coordenações complexas entre nós. Da mesma forma, para eliminar a redundância do tráfego um salto em redes de malha sem fio, um nó deve coletar tal informações de cache de todos os outros nós dos quais o tráfego pode ser ouvido pelo nó do próximo salto.

Em vez de usar uma estimativa de probabilidade de papel aborda o desafio do conteúdo sem fio que ouve demais para RE de camada IP usando previsões. Consequentemente, propomos Eliminação de redundância baseada em previsão com conteúdo Over-audição (PRECO). Em PRECO, o receptor armazena o recebido e o fluxo de dados ouvido em uma cadeia de pedaços. Ele compara os pedaços do pacote recebido com as cadeias cache. Em uma partida, espera-se que a futura entrada é muito provável que os dados correspondam aos pedaços previamente armazenados na corrente. O destinatário envia para

os futuros dados do remetente previsões que incluem os hashes de pedaços na corrente. o o remetente remove o pedaço do pacote de saída se encontrar que seu hash combina com uma previsão. Desta forma, PRECO não depende da estimativa de probabilidade de overhearing; ele remove um pedaço duplicado somente se o receptor tiver armazenado em cache o pedaço. Ele garante um sistema eficaz e robusto de resposta de conteúdo RE sobre links sem fio. Também pode ser usado diretamente em múltiplos Infra-estrutura de AP e redes de malha sem ouvir notificações e comunicação de conteúdo em cache.

O PRECO tem desafios fundamentalmente diferentes da previ- A abordagem de RE baseada em Previsão chamada PACK [19] proposta para o ambiente em nuvem. PACK usa um tamanho grande 8KB o que faz com que ele não consiga identificar o conteúdo de granularidade mais fina redundância. Também leva a um conteúdo ineficaz que ouve demais uma vez que muitos nós podem não ouvir tais grandes pedaços em cheio no ambiente sem fio. Em contraste, o PRECO pretende identificar pedaços duplicados de centenas de bytes em um sub-pacote nivelar e trabalhar em links sem fio de baixa largura de banda. Assim, o O custo de transmissão das previsões deve ser considerado em ordem para perceber os benefícios gerais da RE. Outra questão nova em PRECO é a possibilidade de falta de dados em cadeias de partes. PACK faz previsões e combina não apenas com o recebido Fluxos TCP, mas também sobre os fluxos de dados ouvidos. Assim sendo, PRECO deve ser capaz de identificar o fluxo entre os múltiplos Fluxos contendo partes correspondentes que levam a uma precisão mais precisa previsão, apesar da falta de dados em alguns dados gerais fluxos.

Além disso, exploramos os benefícios da implantação de sub- nível de pacote RE como um serviço de camada IP primitivo em todos os nós redes de malhas sem fio. Essa implantação em toda a rede pode fornecer benefícios de desempenho ao eliminar a redundância Todos os links. Graças ao PRECO que permite uma rede eficiente em toda a rede implantação em um ambiente sem fio, propomos ainda roteamento de fonte ciente de redundância em redes de malha sem fio em obter ganhos de desempenho maiores em toda a rede RÉ. Apresentamos uma "transmissão estimada consciente de redundância tempo "(RETT) métrica. RETT prevê a quantidade total de tempo seria necessário enviar um pacote de dados ao longo de uma rota tomando em relação ao nível de link RE. Ao encaminhar o tráfego do Internet para a rede de malha, um gateway escolhe a rota com o RETT mais baixo.

Em resumo, a contribuição deste trabalho é a seguinte.

1. PRECO oferece uma solução efetiva, eficiente e escalável para a camada de IP baseada na camada de IP de conteúdo overhearing over over wireless links.
2. Um algoritmo de roteamento consciente de redundância é proposto para fur- explorar os benefícios da implantação de RE em toda a rede em redes de malhas sem fio.

II revisa os esquemas existentes para a eliminação de redundância em redes com fio e sem fio. Na seção III e IV, apresentamos o design do PRECO e o algoritmo de roteamento com redundância m, respectivamente. A seção V apresenta avaliação de desempenho. A seção VI conclui este artigo.

2. Trabalho relatado

2.1. Eliminação de redundância de tráfego em redes com fio

Várias técnicas de eliminação de redundância de tráfego (TRE) foram propostos para redes com fio nos últimos anos. Uma abordagem TRE independente do protocolo foi proposta pela primeira vez em [20], que identifica pedaços duplicados no nível de sub-pacote com hash baseado em conteúdo. Vários vendedores comerciais têm desenvolvido tais técnicas em sua "otimização WAN" middle-boxes [21], [22], [23]. A implantação bem-sucedida de As soluções TRE em redes empresariais motivaram a exploração de implantação da TRE em toda a Internet e redundância - roteamento consciente foi proposto para melhorar ainda mais os benefícios de TRE de toda a rede [15], [18]. EndRE foi proposto [17] para eliminando a redundância do tráfego do servidor para o cliente. Há O método usado em [16] elimina a redundância do tráfego com o abordagem de compressão delta. Todas essas abordagens acima faça uso de caches bem sincronizados ou a mesma referência arquivo entre o remetente eo receptor, para que eles não possam ser usados com overhearing para RE em redes sem fio. PACK [19] foi proposto para o ambiente da nuvem [24]. Ele apresenta RE baseado em previsão, em que o receptor compara o dados recebidos com cadeias de partes previamente recebidas e notifica o remetente os pedaços correspondentes. Desta forma, PACK não requer sincronização de cache entre o remetente e receptor. No entanto, o PACK usa um tamanho grande 8KB. Isso faz com que ele não identifique o conteúdo de granularidade mais fina redundância. Também leva a um conteúdo ineficaz que ouve demais uma vez que muitos nós podem não ouvir grandes pedaços na íntegra o ambiente sem fio. Yu et al. [25] cooperativa proposta eliminação de redundância de tráfego de ponta a ponta em ambos os lados da o remetente e receptor para grandes e pequenos pedaços para reduzindo o custo da largura de banda da nuvem.

2.2. TRE em redes sem fio

Em Cache Assimétrico (AC) [11], o receptor encontra em seu cache o segmento de fluxo combinado que tem o máximo número de partes correspondentes em comparação com o tráfego em curso fluxo recebido, e envia de volta seus pedaços de pedaços. O remetente em seguida, executa operações RE com base no seu cache de comentários armazenando os hashes recebidos e o cache normal. SmartEye [26] agrega imagens semelhantes ao mesmo grupo através de uma semântica hashing na nuvem e, em seguida, eliminar a redundância em a transmissão de dados para imagens coletadas do smart termina como imagens existentes na nuvem. Contudo, esses métodos não conseguem alavancar a audição excessiva. Vários métodos foram propostos para alavancar o overhearing sem fio para eliminar transmissões de dados redundantes. No RTS-id [9], o receptor armazena os pacotes ouvidos e o remetente adiciona um pacote especial ID para o pacote 802.11 RTS para que o receptor possa verificar se o pacote de dados a ser transmitido está em seu cache. Ditto [10] and REfactor [12] são duas abordagens RE baseadas em conteúdo. Em Ditto, os roteadores de malha sem fio reassembled os fluxos TCP tocados o servidor para o cliente para reconstruir os pedaços de dados do aplicativo de tamanho aproximadamente 8KB, e armazenar pedaços em seus caches.

Evita transferências de dados ao atender os pedidos de clientes de os roteadores de malha em vez do servidor. REfactor explora mais fino- redundância de granularidade no nível de sub-pacote por camada IP RE com conteúdo que ouve demais. O remetente

estima a sobreposição probabilidade de dados para o receptor, e remove duplicata pedaços apenas se a redução esperada do tempo de transmissão resultando da redundancy removal excede algum limite. No entanto, o ganho de desempenho do REfactor é vulnerável a estimativa imprecisa da probabilidade de overhearing, e também incorre em custos significativos de comunicação ao ser estendido para múltiplas infra-estruturas AP e redes de malha.

3. RE com predição com Overhearing

3.1. Visão geral e um exemplo

Em PRECO, um nó sem fio analisa a carga útil de um recebido ou salva o pacote em pedaços e calcula o valor de hash por cada pedaço. Os pedaços do mesmo fluxo de dados são ligados sequencialmente a uma corrente e armazenados com seu hash valores em um cache de bloco local (Figura 1). Quando um receptor recebe ou ouve um pacote de um remetente, ele compara os pedaços do pacote para o cache local. Se um O pedaço de correspondência é encontrado, o receptor recupera a sequência de pedaços subsequentes após o pedaço de correspondência e envia seus hashes para o remetente como uma previsão. O remetente executa a fragmentação na carga útil do extrovertido pacotes e combina os pedaços contra as previsões de o receptor. Uma vez que encontre uma partida, o remetente remove o pedaço e substitui-lo por um item contendo uma identificação de previsão para confirmar a correspondente previsão de partes. O receptor em seguida, substitui o calção com o correspondente pedaço em sua cache do bloco local. PRECO é vantajoso do que o anterior RE aborda porque os pedaços previstos estão garantidos no cache local do destinatário, o que aumenta a precisão de previsão e, portanto, o rendimento da transmissão. Além disso, PRECO traz benefícios e eficiência em vários AP infra-estrutura, conforme explicado abaixo.

Na Figura 2, o cliente C1 pode ouvir a transmissão de dados da AP2 ao Cliente C2. Para pedaços sucessivos a, b, c, d e e de várias cargas de pacotes em um fluxo de dados transferido de AP2 a C2, C1 com sucesso os pedaços ouvidos a, c, d e E, e colocou-os em uma corrente com seus hashes H a, H c, H d e ele . Observe que o pedaço b está faltando devido à probabilística barulho sem fio. AP1 mais tarde teve um objeto de dados semelhante com os pedaços a, b, c, d e f para enviar para C1. Quando o C1 recebeu Chunk a, descobriu que tinha um pedaço de correspondência na corrente, e depois enviou para AP1 as previsões do bloco que incluem hashes dos pedaços subsequentes c, d e e na corrente. No extrovertido dados, a AP1 descobriu que os pedaços c e d têm hashes correspondentes

com as previsões, e as substituiu pelas calças "I c" e "eu d". C1 pode reconstruir o pacote completo usando o cache pedaços correspondentes às previsões. Como os tamanhos de A predição e a diferença são muito menores do que o tamanho do bloco, PRECO reduz o número de bytes transferidos no ar e melhora rendimento global da rede. Em contraste, REfactor [12] requer comunicação de notificação auditiva entre AP1 e AP2. O AP1 precisa obter periodicamente o conteúdo do cache do AP2 para C1, que contém entradas e estimativas de audição excessiva probabilidades de todos os pedaços presumivelmente ouvidos por C1, que custou um monte de sobrecarga de rede. PRECO precisa resolver os seguintes problemas:

1. • Como realizar a fragmentação e armazenamento em cache para uma redundância de dados na transmissão de dados? (Seção III-B)
2. • Como escolher um fluxo e um pedaço de correspondência para prever para aumentar a precisão da previsão? (Seção III-C)

3.2. Chunking e Caching

1. Chunking Algorithm

- PRECO divide a carga útil de um pacote em pedaços por um algoritmo de fragmentação baseada em conteúdo. Este algoritmo determina os limites do bloco usando o conteúdo em vez de deslocamento, então alterações localizadas no fluxo de dados apenas afetam os pedaços que estão perto das mudanças, o que permite uma robusta identificação de conteúdo duplicado em diferentes objetos de dados. Um número de algoritmos de fragmentação baseada em conteúdo foram propostos, incluindo as impressões digitais de Rabin [20], [10], MAXP [14], SAMPLEBYTE [17] e rolamento baseado em XOR hash [19]. Como REfactor [12], PRECO usa MAXP [14] para definir limites de partes, porque o MAXP fornece uniformemente limites de partes distribuídas em toda a carga útil e impõe um limite inferior no comprimento do pedaço e baixa carga computacional. MAXP seleciona uma posição como limite de pedaço se seu valor de hash for o máximo (ou mínimo) sobre todos os hashes calculados sobre a região de byte- p centrada nessa posição. A carga útil do pacote é dividida em pedaços por esses limites, como mostrado na Figura 1. O tamanho esperado do pedaço é p e todos os pedaços devem ter comprimento pelo menos $p / 2$, exceto o último no final da carga útil [27]. Ignoramos o último pedaço se seu tamanho for inferior a $p / 2$, tal como o pedaço que segue Chunk2 na Figura 1. p deveria ser significativamente maior do que o tamanho da soma de um pedaço de hash e shim em PRECO porque a economia efetiva de largura de banda resultou de uma previsão bem sucedida do bloco é o tamanho do bloco menos o tamanho da soma. Adicionalmente, também limitamos o comprimento máximo de um pedaço para bytes B_{max} . A determinação do tamanho do bloco deve considerar o trade-off entre despesas de previsão e economias de largura de banda. Um tamanho de bloco maior reduz o número de mensagens de previsão, enquanto um tamanho de pedaço menor pode aumentar a eficiência de detecção de bytes redundantes.

2. Caching Received and Overheard Chunks

- Em PRECO, os nós sem fio ouvem as transmissões de fluxos TCP. PRECO identifica cada fluxo por $(src, dst, src\ port, dst\ port)$ tupla, referida como ID do fluxo. Um nó sem fio mantém uma lista de fluxo para gravar as IDs dos fluxos do TCP, que atualmente está recebendo e ouvindo demais. Quando um novo pacote chega, o nó verifica a lista para determinar se o pacote é de um novo fluxo ou de um fluxo existente. Se o pacote é de um fluxo existente, seus pedaços estão ligados a pedaços anteriormente ouvidos no mesmo fluxo. De outra forma, o nó cria uma entrada para o novo fluxo na lista e armazena os cachimbos do pacote. Um período de tempo limite é especificado para cada entrada, e é reiniciado sempre que um novo pacote é recebido do fluxo dessa entrada. Garante o local de tempo de transmissão de pedaços na mesma cadeia. A entrada é excluída se nenhum pacote do fluxo correspondente são ouvidos ou recebidos dentro do período de tempo limite. Lembre-se de que um fluxo consiste de pacotes, um

pacote consiste de pedaços, e um pedaço consiste em bytes de dados. Os sem fio nodos armazenam os pedaços ouvidos / recebidos em seu pedaço local cache com os hashes e números de sequência do pedaços. O número de sequência de um pedaço é o número de sequência de seu pacote TCP contendo mais o deslocamento do pedaço em a carga útil do pacote. Os pedaços de pacotes do mesmo TCP O fluxo está ligado a uma cadeia na ordem de sua sequência números, como mostrado na Figura 1. Além disso, qualquer byte no pedaço tem um número de sequência que é o número de sequência do bloco mais o deslocamento do byte no pedaço. Os pedaços no A cadeia pode não ser dados consecutivos no fluxo, devido a falta de dados devido à audição probabilística e ignorada pedaços pelo algoritmo de fragmentação. Ainda assim, o nó em um multi- A rede sem fio pode ter a chance de preencher as falhas dados incorridos pela audição probabilística, uma vez que pode ouvir o mesmo pacote em diferentes lúpulos.

Os nós sem fio podem receber / ouvir o mesmo pacote várias vezes, quer por mecanismos de retransmissão em 802.11 protocolos MAC e TCP, ou porque os nós em redes de múltiplos saltos pode ouvir a transmissão de um pacote em diferentes lúpulos. Para detectar pacotes duplicados, PRECO armazena o número de sequência TCP de um pacote, juntamente com o seu pedaços no cache. PRECO considera um pacote duplicado se repetiu o número de sequência TCP no fluxo. De Ignorando pacotes duplicados, PRECO garante que apenas um único A cópia do pacote é cortada e armazenada, evitando assim custo desnecessário de processamento e armazenamento.

3.3. Eliminação de Redundância Baseada em Previsão

1. Prediction Algorithm

Ao receber um novo pacote do remetente, o receptor executa o corte e calcula o hash para cada pedaço na carga útil, e então olha para cima esses hashes em seu cache local. Se um hash de um pedaço for encontrado, significa que existe um par duplicado no cache. Com base nas cadeias que contêm os pedaços correspondentes, o receptor encontra os pedaços que provavelmente apareceriam em os próximos dados recebidos, e envia ao remetente os seus hashes como previsões. Abaixo, explicamos como encontrar os pedaços como previsões. A abordagem anterior PACK [19] faz previsões para cada pedaço de segmento ingle porque novos pedaços chegam sequencialmente em Transmissão TCP. Toda vez que um pedaço de correspondência é encontrado, ele re- lança uma sequência de partes subsequentes ao pedaço de correspondência na cadeia para previsões. Tal algoritmo de previsão, como, nunca, é eficiente para PRECO. Isso ocorre porque em PRECO, um pacote pode ter vários pedaços que têm duplicações em o cache. Se a previsão for feita para cada correspondência, múltiplas sequências de pedaços, cada uma seguindo uma correspondência pedaço, seria recuperado para prever a mesma entrada dados, que incorrem em transmissões de previsão extras. Além disso, esses pedaços de correspondência podem ser espalhados por diferentes cadeias, mas algumas cadeias podem não ter muito em comum com a objeto de dados na transmissão e previsões dessas cadeias são inúteis. Para resolver esses problemas, PRECO determina um pedaço de correspondência, o que leva à sequência de partes mais provável que apareça no futuro dados recebidos, e usa o Sequência de pedaços seguindo-o na cadeia para predição.

Nos referimos a um tal pedaço de correspondência como uma âncora de previsão. Para determinar a melhor âncora de previsão, primeiro encontramos a cadeia que tem "correspondência máxima" com os pedaços recebidos e depois decidir qual pedaço de correspondência é usado como o âncora de previsão. Para esse fim, um método direto é escolher a cadeia que contém o maior tamanho total de pedaços de correspondência, e use o maior entre todos os trocados correspondentes naquela corrente como âncora de previsão. Dentro detalhe, denotar as correntes por L_1, \dots, L_k . Suponha que cada cadeia L_i tem pedaços de correspondência $C_{i,1}, C_{i,2}, \dots, C_{i,l_i}$. Deixe $C_{i,m}$ seja do tamanho do pedaço C_{im} . O tamanho total dos pedaços de correspondência na cadeia L_i eu sou. Este método, No entanto, pode não gerar previsões eficientes, pois não considere as distâncias entre os pedaços de correspondência em Transmissão TCP. A eficácia do RE baseado em previsão vem da continuidade do conteúdo duplicado. Se a correspondência os pedaços estão vagamente espalhados na corrente com grandes lacunas entre si, a corrente não está bem combinada com a objeto de dados na transmissão e pode fornecer previsões ruins. Assim, apresentamos uma abordagem de fragmentação para encontrar o Cadeia de "correspondência máxima" com a consideração da distâncias entre os pedaços de correspondência.

2. Transmissão de Previsão e Decodificação de Shim

Para o eficácia das previsões, PRECO limita a faixa de dados previsto a partir de uma âncora de previsão, estabelecendo uma previsão Janela W . Na sequência do bloco seguindo a previsão âncora, um pedaço só pode ser usado para predição se o A distância entre ele e a âncora de previsão é menor que W . O receptor recupera os hashes do bloco dentro da previsão janela, e usa-os para previsões de bits. Cada pedaço Também é atribuído com um ID de previsão que é a sua sequência número na janela. Uma previsão de bloco inclui um hash para o pedaço de entrada esperado. O receptor envia o pedaço previsões para o remetente em uma mensagem de previsão. O pedaço as previsões na mesma mensagem de previsão formam uma sequência. O receptor também mantém um cache de previsão para armazenar previsões enviadas recentemente ao remetente. Ao receber uma mensagem de previsão do receptor, o o remetente extrai as previsões do bloco e os armazena em um cache de previsão local. Começa a comparar as previsões com seus dados de saída. Para cada pacote de saída, o remetente executa o algoritmo de fragmentação (na Seção III-B1) para dividir a carga útil em pedaços e tenta combinar as previsões com esses pedaços. Se achar que um pedaço coincide com um previsão, o remetente substitui o conteúdo e insere um calço em vez disso, incluindo o deslocamento do pedaço no pacote e o número de sequência do pedaço no mensagem de previsão. Depois de receber um pacote contendo um calço do remetente, o o receptor encontra o pedaço correspondente ao número de sequência no calço, e reconstrói o pacote completo, substituindo o calce com o pedaço.

3.4. Redundância-Aware Source Routing In Wireless Mesh Redes

Em redes de malha sem fio, cada roteador de malha compartilha In- acesso de ternet comunicando-se com alguns nós de gateway e fornece conectividade com a Internet para clientes móveis. É muito prometendo melhorar o rendimento da malha sem fio rede-funciona eliminando a redundância no tráfego transferido dos gateways para roteadores de malha. Nesta seção, consideramos para implantar PRECO como um serviço primitivo

de camada IP em todos os nós em redes de malha sem fio e aproveitar os benefícios do roteamento para maximizar a oportunidade de reduzir o conteúdo redundante. Figure 3 gives an example Gateway to show the benefit of the redundancy-aware source routing. Each link is labeled with ETT=3 ETT=2 its ETT (Expected Transmission Time) metric [28] defined A1 as the expected MAC layer duration for a successful transmission. Client C1 requests a data object D1 consisting of chunks a, b, c and d, and the gateway chooses the route Gateway-A1 - A3 to transfer the data object. Traditional ETT-based source routing protocol chooses the route with the minimum sum of ETT of links on the path. Thus, when C2 requests a data object D2 consisting of chunks a, b, c and e, the gateway chooses the route Gateway- A2 - A4 - A5 with ETT metric of 8. Suppose each router on a path caches the transferred data object and is capable of removing redundant content. Then, if the route Gateway - A1 - A3 - A5 is used for transferring D2, the redundant content a, b, c can be removed over links Gateway-A and A1 - A, and only e is required to transfer over these links. Suppose the size of each chunk is equal to a packet size, then according to the computation of ETT, the time required to transfer D2 over link Gateway-A1 and A1 - A3 is 3 and 4 respectively. The total ETT for D2 is $3+4+16=23$. The traditional ETT- based route Gateway-A2 -A4 - A5, however, has a total ETT for D2 is $8+12+12=32$.

1. Métrica de Roteamento com Redundância

Propomos uma "transmissão estimada consciente de redundância Time" (RETT) de um link derivado do ETT metric [28]. ETT é calculado por $ETT = ETX \times SB$ onde ETX [29] é Transmissão esperada Contagem que estima o número de Retransmissões necessárias para entregar um pacote sobre o link, S é o tamanho médio do pacote e B é a largura de banda. Suponha que um pacote tem uma razão de redundância média em comparação com o cache de pacotes em um link. Então, o tempo de transmissão esperado para o pacote com a consideração da redundância A eliminação sobre o link, ou seja, RETT, pode ser calculada por Esta métrica de link pode refletir com mais precisão a transmissão Tempo para transmissão de conteúdo aproveitando o redundante conteúdo nos roteadores. No entanto, também tem problemas. Primeiro, ignora a sobrecarga das transmissões de previsão. Em segundo lugar, a razão de redundância a varia entre os pacotes e depende do conteúdo dos pacotes. Em terceiro lugar, é extremamente caro para um protocolo de roteamento para calcular uma rota ótima para cada pacote individual. Para lidar com esses problemas, nossa solução é considerar o roteamento ideal para um fluxo TCP completo ou grande dados em massa no fluxo, de modo que uma rota ideal possa ser decidida por um grupo de pacotes com maior probabilidade de ter o perfil de redundância semelhante. Para a transmissão de previsão sobrecarga, uma vez que uma mensagem de previsão contém várias partes previsões. A sobrecarga de uma mensagem de previsão é amortizada sobre todas as predições de pedaços na mensagem. Além disso, cada um a predição de bloco tem um tamanho muito menor (por exemplo, 32 bits hash + ID de previsão de 10 bits) do que o tamanho médio do pedaço (por exemplo, 256 bytes). Portanto, a sobrecarga de previsão é insignificante, especialmente quando os dados tiverem redundância suficiente. A métrica RETT de uma rota é a soma das métricas do link. Prevê o total quantidade de tempo que demoraria para enviar os dados em

massa ao longo de um rota, levando em consideração o nível de link RE.

2. Routing Protocol

Com base nas discussões acima, propomos uma redundância - protocolo de roteamento de código ciente para um gateway para encaminhar tráfego da Internet para a rede de malha. Para a computação RETT, podemos usar o método no anterior trabalho [29] para calcular o ETX de cada link. Cada roteador de malha transmite sondas de link de tamanho fixo em um período médio, e calcula a taxa de entrega contando o número de sondas recebido durante uma janela de tempo. Então, os roteadores relatam o estatísticas de volta para o gateway. Para obter a razão de redundância Para um dado em massa, o gateway armazena todos os dados a granel transferidos juntamente com suas informações de caminho de roteamento. Usando o anteriormente algoritmo de fragmentação introduzido, o gateway também divide cada transferiu dados em massa para pequenos pedaços, calcula o hash por cada pedaço, e armazena-os no seu cache. Quando encaminhando um novo volume de dados, o gateway divide-o em pedaços e olha para cima seus hashes no cache do pedaço para encontrar o volume - dados no cache que tem o maior número de duplicatas pedaços com os novos dados em massa, denotados por N_{max} . Suponha que o novos dados em massa têm N número de pedaços. Então, a redundância A relação dos novos dados em massa é $a = N_{max} / N$. Suponha que a rota para os dados em massa que tem N_{max} O número de partes duplicadas consiste em nós n_1, n_2, \dots, n_k . Com PRECO que executa RE com base na previsão do receptor, qualquer link com n_i como o nó receptor pode possivelmente obter o benefício da RE. Para garantir uma transmissão de previsão negligenciável custo comparado com a redundância a ser removida, podemos Dê um limiar a T . Ou seja, se N_{max} / N maior a T , a métrica RETT é calculado pela Equação (2), caso contrário, é igual a ETT. Depois de calculando a métrica RETT de todos os links, o gateway é executado O algoritmo de Dijkstra para encontrar a rota com o RETT mais baixo. No acima, o cálculo da métrica RETT do link em O protocolo de roteamento proposto não considera o efeito de conteúdo ouvido para redundância de eliminação de PRECO. Considera apenas o efeito de dados armazenados em cache no caminho transferir. Alguns links podem não ter nós no caminho, mas eles tem nós que ouviu o conteúdo duplicado. Com estas overhearing nodos como receptores de nós em links e redundância serviço de eliminação, esses links podem ter menor expectativa tempo de transmissão do que o ETT, e, portanto, pode ser usado para encontre uma rota melhor. No entanto, para calcular RETT para estes links, o gateway precisa ter uma estimativa precisa do probabilidade de overhearing de nós, que foi mostrado para seja difícil. Devido a isso, nosso protocolo de roteamento calcula as rotas com base em que o gateway possui o roteamento definitivo informações de dados em massa transferidos anteriormente. Enquanto isso, o benefício do conteúdo de audição pode ser oportunista explorado como apresentado anteriormente no nosso método RE. que é, cada receptor prevê os dados que serão transmitidos para baseia-se nos dados recebidos e ouvidos em seu cache para eliminação de redundância na transmissão de dados.

3.5. Performance Evaluation

Neste artigo, desenvolvemos um simulador usando Java para realizar experimentos de simulação de rastreamento real para avaliar o desempenho de PRECO. Nós reunimos

traços de dados sem fio de executar o aplicativo do YouTube em dois smartphones, que são iPhone 6 plus e Xiao Mi 3. Para capturar o tráfego, nós deixamos o tráfego passar por um laptop (Lenovo T420 com Windows 10) e usamos o software Wireshark para capturá-lo. Especificamente, nós conectamos o laptop à Internet com o cabo e configuramos o modo Hotspot no laptop. Os dois smartphones estão conectados com o laptop usando o WiFi compartilhado pelo laptop. Como resultado, os dois smartphones podem acessar a Internet. Os dois smartphones assistiram o mesmo selecionado aleatoriamente vídeos durante 20 minutos, e depois assistimos dois vídeos diferentes com conteúdo similar durante 10 minutos, como mostrado na Figura 4. Esse processo foi repetido duas vezes. Nós coletamos dados 60 minutos a dia por 7 dias no total. O rastreamento de dados em direção ao iPhone 6 plus é denotado como T raça 1 e o outro é denunciado como T corrida 2. Finalmente, obtivemos dados de 1.71GB para T race 1 e 1.64GB para T corrida 2. Comparamos PRECO com EndRE [17], assimétrico Caching (AC) [11] e REfactor [12] usando as seguintes Métricas:

1. Eficiência RE. É a proporção dos bytes totais de redução de dados redundantes para o volume total de dados e calculados por $(V_{No-RE} - V_{RE}) / V_{No-RE}$, onde V_{No-RE} é o total de dados volume e V_{RE} é o volume total dos dados reduzidos. Uma maior eficiência de RE significa uma maior taxa de previsão, eventualmente, uma maior precisão de previsão.
2. • Sobrecarga da rede. É calculado por $V_{Overhead} / V_{dados}$, onde $V_{Overhead}$ é o volume de dados das mensagens de previsão e hashes na transmissão de conteúdo e os dados V são do tamanho do total de dados de conteúdo transmitido.
3. • Relação de economia de banda. É calculado por $(V_{No-RE} - V_{RE} - V_{Overhead}) / V_{No-RE}$ para mostrar a largura de banda final poupança causada por RE e rede de redução de despesas gerais.

(a) Configuração de Simulação

Avaliamos a eficiência do PRECO com três diferentes cenários de simulação. Primeiro, implantamos uma infra-estrutura AP com um cliente associado em nosso simulador para avaliar o desempenho de PRECO sem conteúdo ouvido. Então, nós simulamos uma rede pequena com dois APs, cada um com um cliente associado, para avaliar o desempenho de PRECO usando Conteúdo que ouve demais. Finalmente, construímos um tipo de rede sem fio rede de malha para avaliar os benefícios da redundância encaminhamento. PRECO usa MAXP para encontrar os limites do pedaço e o traço é dividido em pedaços por esses limites. Limitamos os tamanhos do bloco a [256,1024] bytes. Nós estabelecemos o limiar de distância de fusão $d_T = 1500$ bytes e a inicial Tamanho da janela de previsão $W_0 = 4KB$.

(b) Point-to-Point RE Efficiency

Nesta simulação, o nodo AP transfere sucessivas T corrida 1 e T corrida 2. PRECO permite ao cliente utilizar T corrida 1 no cache para fazer previsões quando o AP começou a transmitir T corrida 2. O AP remove bytes redundantes com base em as previsões do cliente. A Tabela I mostra experiências resultados para as diferentes métricas sem eliminação de redundância ("Não-RE") e com PRECO. Como podemos ver, para a corrida T 2, PRECO reduz o tráfego volume de 1,64 GB a 0,51 GB. Além disso, quando a transferência-anel T raça 1, há também considerável redundância, mais de 16 detectado por PRECO mesmo sem transmissão de

dados anterior. Isso indica que tal redundância existe na própria corrida T 1. Em comparação com a significativa redução de tráfego, a rede a sobrecarga é insignificante. Podemos ver que o PRECO é efetivo em reduzindo o custo da largura de banda tanto para a corrida T quanto para a corrida T 2. Mais distante- mais, ao transferir T raça 2, PRECO pode alcançar maior relação de economia de banda, em comparação com a transmissão T race 1. Isso ocorre porque durante a transmissão de T race 2, PRECO pode reduzir os dados redundantes, não apenas usando o seu próprio anterior tráfego, mas também usando o tráfego anterior da corrida T 1.

(c) Benefícios do conteúdo Overhearing

Nesta simulação, dois nós AP, doados pela AP1 e AP2, e dois clientes, designados por C 1 e C 2, são implantados. AP 1 transfere a corrida T 1 para C 1 e C 2 funciona em modo promíscuo para que possa ouvir a transmissão entre AP1 e C 1. Ao mesmo tempo, AP2 transfere T corrida 2 para C 2. Medimos a eficiência de RE, a sobrecarga de rede e a economia de banda razão de PRECO, EndRE, AC e REfactor com cache diferente tamanhos e probabilidades de overhearing. Salvo indicação em contrário, a probabilidade de overhearing foi ajustada para 800 tamanho do cache foi definido como 200MB.

A Figura 5 (a) e 5 (b) mostram, respectivamente, a eficiência RE de diferentes métodos com o tamanho do cache no receptor variando de 20MB para 200MB, e que com o overheair mudança de probabilidade de 0A eficiência RE segue PRECO, REfactor, EndRE, AC. EndRE e AC geram menor eficiência de RE porque não fornecem porto ouvindo, e eles perdem as oportunidades de eliminar redundância baseada no tráfego ouvido. Além de usar Overhearing, PRECO pode fazer uma previsão mais precisa sobre dados redundantes do que AC. AC escolhe a mensagem de feedback em cache do receptor ao encontrar o segmento de fluxo combinado com o número máximo de partes correspondentes em comparação com o fluxo recebido. O algoritmo de predição de PRECO usado no O receptor considera não apenas o número de partes correspondentes, mas também a distância entre eles, que fornece mais precisa previsão do que AC. Observamos ainda isso, embora REfactor também suporta overhearing, produz eficiência de RE menor do que PRECO. REfactor toma a decisão da remoção de partes com base na estimativa de probabilidade de overhearing, que pode não seja preciso porque o resultado da estimativa é vulnerável a as mudanças dinâmicas na rede. Se a estimativa for não é preciso, o REfactor pode tomar uma decisão errada sobre eliminação de redundância.

Da Figura 5 (a), também podemos encontrar isso com o aumento de o tamanho do cache, a eficiência RE aumenta. A razão é que O cache com um tamanho maior armazena mais cachimbos, o que melhorar a probabilidade de eliminação de redundância e, em seguida, Eficiência RE. Além disso, vemos isso, para PRECO e REfactor, o A eficiência de RE aumenta com a probabilidade de overhearing. Isto é porque essa maior probabilidade de overhearing traz muito mais pedaços redundantes de outras transmissões, o que melhora a precisão da predição e, em seguida, a eficiência RE. A

Figura 6 (a) e a Figura 6 (b) mostram, respectivamente, a rede sobrecarga de diferentes métodos com tamanho de cache variando de 20MB a 200MB e que com probabilidade de overhearing de 0a maior sobrecarga da rede entre todos esses métodos RE. A razão é que no REfactor, para estimar a probabilidade de overhearing para dados transmitidos, dois APs precisam comunicar uns com os outros, o que incorre em uma grande quantidade de sobrecarga de rede. Para o EndRE, a sobrecarga da rede vem dos hashes da peça na transmissão de conteúdo. Em AC, a sobrecarga de rede inclui os maus-tratos no feedback mensagens e transmissão de conteúdo. Para PRECO, a rede a sobrecarga inclui os hashes de bloco e os IDs nas mesas de previsão. sages e IDs na transmissão de conteúdo. Como resultado, EndRE, AC e PRECO geram muito mais sobrecarga de rede do que o REfactor. Da Figura 6 (a), também encontramos que com o tamanho do cache aumentado As despesas gerais da rede aumentam. Para AC e PRECO, o O motivo é que um cache com um tamanho maior trará mais dados redução e leva a mais hash sobrecarga na transmissão. Para REfactor, à medida que o tamanho do cache aumenta, o remetente armazenará em cache mais pedaços e produzir mais sobrecarga para ouvir estimativa de probabilidade. Em EndRE, para descarregar o hash informando do receptor para o remetente, as transferências do remetente os maços de dados de todos os conteúdos, independentemente da os pedaços de dados são redundantes ou não, então a carga de hash continua constante. A Figura 6 (b) mostra que as sobrecargas da rede em PRECO e REfactor aumentam com o aumento da sobre- probabilidade auditiva. A maior probabilidade de overhearing produz maior eficiência de RE, o que leva a mais hashes transmitidos devido à redução de dados. Para EndRE e AC, o aumento da probabilidade de overhearing não tem efeito em sua rede overheads porque eles não consideram o conteúdo ouvir demais. A Figura 7 (a) e a Figura 7 (b) mostram, respectivamente, a largura de banda relação de economia de diferentes métodos com o tamanho do cache no receptor variando de 20MB para 200MB, e que com o overheair mudança de probabilidade de 0A relação de economia de largura segue PRECO, REfactor, EndRE, AC. O resultado é consistente com o da Figura 5 (a) e da Figura 5 (b) uma vez que a economia de largura de banda é causada principalmente pela eficiência RE e as despesas gerais da rede representam apenas uma parte muito pequena do custo da largura de banda.

(d) Benefícios do Roteamento com Redundância

Na simulação, implantamos uma rede de malha com 5 linhas e 5 colunas no total. Nós marcamos o nó como 0, 1, 2, 3, 4), onde m e n são o índice de linha e coluna respectivamente. A 0,0 é o gateway desta rede de malha. o gateway como o nó de origem envia os pacotes para o cliente C 1, associado a A 4,3, cliente C 2, associado a A 4,4, e cliente C 3, associado a A 3,4. Nesse cenário, estabelecemos o overhearing cobertura para 1. Isso significa que apenas o nó do remetente O nó vizinho pode ouvir a transmissão de dados, o que confirma o cenário prático, já que os nós mais próximos têm muito mais chance de ouvir os dados enviados pelo remetente. C 2 pode ouvir a transmissão de dados de A 4,3 para C 1 e A 3,4

para C 3. O gateway envia a corrida T 1 para C 1 e C 3 sucessivamente, Ao mesmo tempo, envia T corrida 2 para C 2. Nós nos preocupamos com a taxa de transferência de rede de A 0,0 a C 2. A taxa média de data para Cada link varia de 800Kps a 1200Kps

Para investigar os benefícios da nossa proposta protocolo de roteamento com redundância, implementamos três roteiros se aproxima. (1) roteamento baseado em ETT: o gateway determina a rota ótima para um receptor usando a métrica ETT, e lá não existe uma implantação PRECO em toda a rede para executar RE todos os links; (2) roteamento com redundância sem conteúdo overhearing: o PRECO de toda a rede é implantado, mas sem O conteúdo em excesso, ou seja, apenas o cache no caminho está habilitado para RE e o gateway calcula a métrica RETT para selecionar a rota ótima para um receptor; (3) encaminhamento com redundância com conteúdo que ouve demais: é como a segunda abordagem, exceto que o overhearing do conteúdo está ativado no PRECO. O cache O tamanho no receptor foi ajustado para 200MB. Nós normalizamos o taxa de transferência (bytes / segundo) do roteamento ciente de redundância por esse do roteamento baseado em ETT. A Figura 8 (a) mostra o rendimento normalizado do gate- caminho para o cliente C 2 quando a probabilidade de overhearing foi variada de 0com o roteamento baseado em ETT, nossa recomendação de redundância proposta o roteamento pode produzir 20Ao contrário do roteamento baseado em ETT, o roteamento com redundância usa RE. Ou seja, o gateway orienta o tráfego através do nós com alta redundância, o que ajuda a reduzir a transmissão tempo e melhorar o rendimento. Quando o conteúdo está ouvindo demais é considerado, o throughput é ainda melhorado. Nós podemos Veja como a probabilidade de overhearing aumenta, o throughput aumento cresce. Quando a probabilidade de overhearing é de 90o throughput é melhorado em mais de 60À medida que aumenta a probabilidade de overhearing, o nó recebe pedaços mais redundantes de outras transmissões, o que melhorar a precisão da predição no PRECO e reduzir ainda mais os dados redundantes. Nós também testamos nosso protocolo de roteamento com redundância usando diferentes métodos RE. O normalizado O throughput entre o gateway e o cliente é plotado na Figura 8 (b).

Vemos isso em comparação com o ETT abordagem de roteamento, tanto as abordagens RE podem melhorar a throughput devido ao roteamento com redundância e RE métodos. Verificamos também que a melhoria da taxa de baixo PRECO_i REfactor_i EndRE_i AC. Isto é porque alto RE eficiência significa menos transmissão de dados em cada link, o que reduz mais tempo de transmissão no caminho total e produz maior melhora na produção. Uma vez que o RE ef a sigilo PRECO_i REfactor_i EndRE_i AC, o PRECO supera outros métodos sobre a melhoria do throughput. o as melhorias de produção crescem em PRECO e REfactor como a probabilidade de overhearing aumenta devido às mesmas razões como na Figura 5 (b).

4. Conclusão

Neste artigo, propusemos uma camada IP baseada em predição RE método com conteúdo ouvido chamado PRECO para wireless redes. Em PRECO, os receptores sem fio comparam as entradas pacotes com pacotes recebidos ou ouvidos anteriores, prever futuros dados recebidos, e enviar seus hashes para o remetentes. Um remetente sem fio remove pedaços de dados redundantes que já existe no cache do receptor, comparando o hashes de pedaços de dados de saída com as previsões de o receptor. Nós também propusemos novos algoritmos de previsão que permitem que PRECO efetivamente melhore a precisão da predição e economia global de largura de banda. PRECO é vantajoso do que Métodos de RE anteriores baseados em overhearing em dois aspectos. Primeiro, por camada de IP baseada em predição RE, PRECO não requer estimativa de probabilidade de overhearing. Em segundo lugar, não precisa comunicação de conteúdo do cache e coordenação complexa a- mong nodos sem fio ao serem implantados em vários AP infra-estruturas e redes de malha. Assim, ele permite eficiência RE de toda a rede com overhearing de conteúdo para rede sem fio trabalho. Exploramos o serviço de camada IP em toda a rede em redes de malha sem fio, e propôs um conhecimento de redundância protocolo de roteamento para melhorar ainda mais seu benefício. Traçada Os resultados da simulação mostram que o PRECO fornece significantes per- benefícios de formatura em comparação com outros métodos RE. Dentro o futuro trabalho, exploraremos como habilitar o gateway para aprender eficientemente os fluxos de dados gerais de todos os nós para determinação da rota em redes de malha.