

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

Apresentação I

Rafael Gonçalves de Oliveira Viana¹

9 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 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. Consequentemente, 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)

Referências