Universidade de Aveiro
2009

Departamento de Electrónica,
Telecomunicações e Informática

Hélio Edgar Nascimento Araújo

Controlo de Mobilidade com Segurança em Redes Estruturadas 802.11

DOCUMENTO PROVISÓRIO

Universidade de Aveiro
2009

Departamento de Electrónica, Telecomunicações e Informática

Hélio Edgar Nascimento Araújo

Controlo de Mobilidade com Segurança em Redes Estruturadas 802.11

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Electrónica e Telecomunicações, realizada sob a orientação científica do Dr. André Zúquete, professor auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro.

O júri

presidente

Doutor José Luís Oliveira
Professor Associado da Universidade de Aveiro

vogais

Doutor Carlos Nuno Ribeiro
Dep. De Engª Informática do IST/UTL, Lisboa

Doutor André Zúquete (Orientador)
Professor Auxiliar da Universidade de Aveiro

Doutor Paulo Jorge Salvador Serra Ferreira (Co-Orientador)
Professor Auxiliar Convidado da Universidade de Aveiro

Agradecimentos

Agradeço o meu orientador de projecto, Dr. André Zúquete, pelo enorme apoio e grande paciência na execução desta dissertação. Ao meu colega Rodolphe Marques que me apoiou durante todo o projecto e sem o qual seria absolutamente impossível fazer seja o que for neste trabalho. Agradeço aos meus colegas que me acompanharam no meu percurso académico. Por fim falta agradecer a minha família, especialmente a Mamã e o Papá, que “bancaram” tudo isto para garantir o meu futuro, e com o seu esforço conseguiram dar-me tudo o que eu sempre precisei. Obrigado.

Resumo

Esta dissertação aborda o problema da gestão da mobilidade com segurança em redes 802.11. Assim, começa por apresentar um estudo detalhado do protocolo 802.11, do *handoff* de dispositivos móveis entre pontos de acesso e de soluções apresentadas por diversos autores com o objectivo de reduzir o tempo dispendido neste processo, com e sem segurança associada. Seguidamente, são apresentadas métricas e atributos de rede que podem ser considerados no estabelecimento de políticas de mobilidade que gerem as transições de AP que cada dispositivo móvel efectua. Uma vez feito este estudo inicial, é apresentada uma solução que potencía *handoff*s rápidos e seguros em redes estruturadas 802.11 e que minimiza o tempo da sua preparação. Este novo protocolo representa uma evolução do trabalho desenvolvido por Rodolphe Marques no trabalho intitulado "Segurança e Mobilidade em Redes Estruturadas 802.11" referenciado em [1]; a sua novidade consiste em usar tramas 802.11 de reconhecimento da rede (***Probe Request*/*Response***) para difundir associações de segurança com os pontos de acesso ao alcance de cada dispositivo móvel. A nova abordagem implica mudanças reduzidas na arquitectura de rede considerada em [1] e permite que, no âmbito das operações de reconhecimento de pontos de acesso, que são comuns e necessárias, um equipamento móvel instale paralelamente associações de segurança nos APs que poderá vir a usar num futuro próximo, ou seja, todos os que estão ao seu alcance.

Abstract

This thesis handles the problem of mobility management with security in 802.11 networks. Therefore it begins by presenting a detailed study of the 802.11 protocol, the handoff process of roaming mobile nodes between access points and solutions presented by many authors with the common goal of reducing the time spent in this process, with and without associated security. After this we present metrics and attributes of the network that may be considered on the establishment of mobility policies that handle the AP transitions made by every mobile node. Once finished this initial study we present a solution that enhances fast and secure *handoffs* in structured 802.11 networks and minimizes the time spent in its setting. This new protocol represents an evolution on the work developed by the author Rodolphe Marques in his work named “Security and Mobility in 802.11 Structured Networks” referred in [1]; its new feature consists in using 802.11 network scanning frames (Probe Request/Response) to distribute security associations to all access points in range of each mobile node. This new approach implies some changes on the architecture proposed in [1] and allows a mobile node to install security associations simultaneously while browsing the neighborhood for access points that may be used in a near future.

Índice

Lista de Figuras

Lista de Tabelas

Capítulo 1

Introdução

A realidade tecnológica actual conta com o protocolo IEEE 802.11 para estabelecer e manter conectividade a uma rede sem fios dispersa a nível mundial. Os pontos de acesso (*Access Points*, APs) podem ser encontrados em grande concentração em locais onde é oferecida mobilidade ao utilizador mantendo as ligações à rede activas dentro da área de abrangência de cada AP. Porém não se pode afirmar que a mobilidade dos nós móveis (*Mobile Nodes*, MNs) seja transparente para o utilizador, de facto demonstra-se ainda insuficiente para suportar serviços exigentes em termos de perdas e atrasos durante o processo de transição entre APs, onde o utilizador experimenta consequências adversas na comunicação.

Em [2], Mishra e Arbaugh apresentam um estudo do processo de *handoff* a nível da camada MAC onde concluem experimentalmente que as latências medidas utilizando os métodos de *handoff* hoje aplicados ultrapassam largamente as exigências requeridas por aplicações sensíveis a atrasos, como por exemplo as aplicações VoIP, onde o atraso máximo recomendado não deverá exceder os 50 milissegundos de forma a manter a comunicação activa. Aplicações VoIP são aqui usadas como referência pois, de entre inúmeras aplicações, estas apresentam maior sensibilidade à latência que qualquer outra.

O atraso experimentado pelos dispositivos móveis durante a transição de AP impossibilita a comunicação contínua e interrupta, o que se traduz na terminação da comunicação dessas aplicações exigentes em termos de atraso, jitter e perdas.

1.1 - Enquadramento

O processo de *handoff* pode ser decomposto em quatro fases distintas. A fase de descoberta, onde o MN efectua a perscrutação do meio em que se encontra inserido em busca de APs a servir a rede. Não está definido o momento em que deve ser levada a cabo esta fase, sendo deixado ao critério do fabricante o momento em que esta deve ter lugar, seja continuamente durante a comunicação, intercalado com a transmissão de dados normal, seja no início do processo de transição, logo que o *handoff* é disparado.

A fase de descoberta é a mais crítica de todo o processo em termos de latência representando 90% do tempo do atraso conferido pelo processo de *handoff*, quando feita aquando a transição, tal como exposto em [2].

O *handoff* compreende a fase de disparo do processo de *handoff*, onde o MN toma consciência da necessidade de se associar a outro AP e despoleta o processo propriamente dito.

Quer seja feita a descoberta do meio proactivamente ou reactivamente durante o processo de *handoff*, é necessária a fase de selecção do AP, onde o MN escolhe, de entre os candidatos encontrados na fase anterior, qual o mais apropriado para o servir.

A fase final denomina-se compromisso. O MN autentica-se com o AP seleccionado na fase precedente, termina a associação actual com o AP que o servia com qualidade degradada e formaliza a associação com o novo AP. Para que isto seja possível é necessário que o MN se autentique com este novo AP e posteriormente se associe. Todas estas fases pertencem tipicamente ao conjunto de operações levadas a cabo pela camada MAC da pilha protocolar.

Adicionalmente poderá ser necessário realizar o handover da camada IP, que pode ser concretizado, por exemplo, pelo MIP (Mobile IP Protocol). A latência típica desta fase de compromisso da camada MAC pode ser considerada diminuta quando tratada nos termos de autenticação/associação 802.11. Contudo, na situação onde são usados protocolos de segurança, o processo toma um rumo mais extenso temporalmente, quando se implementa o 802.11i, usando o protocolo de autenticação 802.1X. Então, na fase de compromisso, além da autenticação/associação 802.11 é ainda feita a autenticação mútua 802.1X entre o MN e o AP, que se demonstra extremamente demorada (aproximadamente 1 segundo, tal como exposto em [3]).

Durante o estudo dos critérios de gestão da mobilidade do MN foram evidenciados dois tipos principais de métodos de ataque ao problema do *handoff*: um grupo de autores empenha-se na redução do impacto que a latência da fase de descoberta do *handoff* tem sobre todo o processo. Outro grupo de autores apresenta soluções que visam a redução da latência adicional conferida pelas configurações e troca de informação dos protocolos de segurança durante o processo de *handoff*.

Um método onde a fase de descoberta seja melhorada irá beneficiar de melhores resultados em termos de atrasos adicionados ao *handoff*. Seguindo um pouco mais além desta observação, se outras fases críticas poderem ser efectuadas proactivamente, em paralelo com a fase de descoberta, sem que daí advenham custos significativos, então todo o processo é melhorado.

O protocolo 802.1X atribui um nível elevado de segurança ao processo de autenticação e controlo de acesso à rede sem fios permitindo a autenticação mútua entre o MN e a rede de acesso gerida por uma organização. Devido ao facto de o standard 802.1X ter sido originalmente desenhado para redes cabladas, o problema de *handoff*s rápidos dos dispositivos móveis em redes sem fios 802.11 com este suporte de segurança não se encontra devidamente desenvolvido de forma a serem observadas latências aceitáveis que permitam uma contínua comunicação do MN com os APs que o servem. No caso particular da segurança, em redes sem fios com suporte de autenticação 802.1X, o tempo de *handoff* durante o qual a comunicação de dados não se processa é extenso demorando cerca de 1 segundo, tal como demonstrado em [3]. O tempo dispendido na autenticação mútua entre o MN e o novo AP irá sempre contribuir para o tempo onde não existe comunicação de dados durante o *handoff*. Tais períodos de tempo são demonstrados na Figura 1.1.

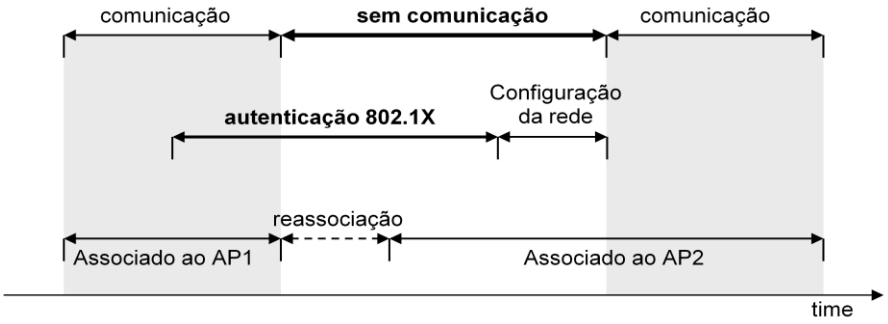

Figura 1.1: Períodos de comunicação e falta de comunicação durante a reassociação com (re)autenticação 802.1X.

1.2 - Objectivo

Evidencia-se o problema que pretendemos resolver com este trabalho: como gerir a mobilidade de dispositivos móveis de forma a minimizar o impacto dos processos de *handoff* (idealmente tornando-os transparentes para o utilizador), para que a qualidade da comunicação seja maximizada.

No fundo pretende-se criar a ilusão de conectividade contínua enquanto o MN se desloca através da rede de acesso, garantindo a melhor qualidade de serviço possível.

O objectivo central deste trabalho foi estudar o processo de *handoff* em redes 802.11 com a finalidade de implementar mecanismos eficazes de gestão da mobilidade do MN tornando o processo de *handoff* menos moroso, maximizando a ilusão de conectividade contínua do dispositivo móvel

com a rede Infra-Estruturada a que se encontra apenso. Neste sentido, foi prestada especial atenção a problemas de segurança, onde se estudaram e conceberam mecanismos que permitem reduzir o tempo gasto durante um *handoff* devido a questões de segurança.

Neste trabalho faz-se um estudo detalhado dos critérios de gestão de mobilidade em redes 802.11 com o objectivo de reduzir a latência total do processo de *handoff*, bem como uma análise detalhada das metodologias e métricas que podem ser manuseadas para atingir este objectivo e os critérios que devem ser avaliados para decidir o momento em que a associação do MN deve ser transferida de AP para que seja providenciada ao utilizador a melhor qualidade de serviço que a rede possa oferecer.

Apresentada esta questão também se propõe uma nova abordagem para tratar reautenticações mútuas intra-domínio de MNs em redes sem fios 802.11 com suporte de segurança 802.1X.

1.3 - Contribuição

A nossa abordagem vai de encontro aos princípios de reassociação de MNs do standard 802.11, ou seja, o MN é autenticado por um AP candidato, enquanto ainda está associado ao AP actualmente a servir o dispositivo, de seguida reassocia-se a esse AP candidato. Com este modelo, o tempo de corte na comunicação durante o *handoff* é limitado ao tempo necessário para realizar a reassociação e configuração da rede, excluindo a latência da (re)autenticação (ver a figura 1.2).

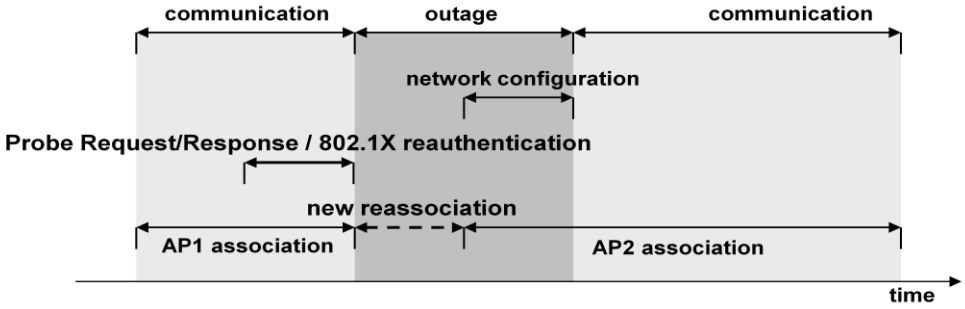

Figura 1.2: Períodos de corte de comunicação durante as reassociações com a nossa nova abordagem.

A solução apresentada permite um paralelismo funcional com os métodos e políticas existentes que confiram ao MN a responsabilidade e controlo da mobilidade, representando um protocolo funcional que melhora a experiência de mobilidade do MN ao resolver o problema da latência do *handoff* de forma optimizada.

Esta abordagem é inovadora pois nunca antes fora usada a fase de perscrutação para efectuar operações de segurança.

Em suma, as nossas contribuições são as seguintes:

- Uma análise detalhada dos mecanismos actuais propostos para redução da latência do processo de *handoff*.

- Um estudo das métricas e métodos disponíveis para implementar políticas de decisões relativas a quando e como realizar o *handoff*.

- Estudo e apresentação de uma abstracção de um protocolo de reautenticação 802.11/1x sem mapeamento em mensagens 802.11 concretas.

- Uma nova abordagem deste protocolo de reautenticação usando a fase de perscrutação, ou seja, com tramas *Probe Request/Response* apresentada no Capítulo 4.

Durante o desenvolvimento deste trabalho foi publicado o artigo "Fast 802.11 Handovers with 802.1X Reauthentications" que apresenta as nossas soluções de implementação do *handoff* rápido com segurança, publicado na revista "Security and Communication Networks - Special Issue, 2009".

1.4 - Estrutura do Documento

Este trabalho é organizado da seguinte forma: No Capítulo 2 é exposta a tecnologia IEEE 802.11, os vários protocolos associados (em especial o IEEE 802.11i) e explicado o processo de *handoff*. No Capítulo 3 apresentam-se os trabalhos relacionados. No Capítulo 4 apresenta-se o protocolo principal no qual nos baseámos na realização deste trabalho [1] e a nossa nova solução para o problema de *handoff* com segurança. Finalmente o Capítulo 5 apresenta as conclusões.

Capítulo 2

Contexto

2.1 - Redes Sem fios IEEE 802.11

Actualmente as redes sem fios encontram-se extremamente difundidas por todo o mundo. A investigação relativa a esta tecnologia teve início nos anos 90. Nessa altura, o padrão 802.11 era visto como uma alternativa promissora opcional ao padrão IEEE 802.3 (Ethernet). As redes cabladas Ethernet interligavam a maioria dos utilizadores empresariais à Internet. Nos primórdios do 802.11 o suporte de aplicações multimédia e voz com necessidades de soluções QoS não era uma característica prioritária do padrão. Contudo, no virar do milénio, demonstrou-se óbvia a necessidade de aplicações com requisitos de QoS. De igual forma, as questões de segurança, que não recebiam a atenção devida, revelam-se cruciais actualmente. Desta forma foram desenvolvidos métodos e protocolos para promoção da qualidade de serviço, autenticação e cifra para garantir autenticidade, privacidade e confidencialidade, integridade da rede e controlo de acesso. As características físicas da camada mais profunda do modelo OSI (camada física) também sofreram avanços que se reflectiram nos aumentos da velocidade de comunicação e área de cobertura de uma célula.

2.1.1 - Arquitectura de uma Rede 802.11

O padrão IEEE 802.11 permite a estruturação de uma rede de duas formas distintas:

• **Ad-Hoc:** nesta organização da rede os nós móveis (*Mobiles Nodes,* MNs) são capazes de comunicar entre si na base de um modelo P2P (*Peer-to-Peer*). Os MNs que constituem uma rede Ad-Hoc formam uma IBSS (*Independent Basic Service Set*). Esta estrutura de rede 802.11 não será abordada neste trabalho.

• **Infra-Estruturada:** Nesta organização de rede os MNs comunicam com pontos de acesso (*Access Points,* APs). O AP funciona como um *switch* (ou bridge) que possui interfaces ou portos para comunicar com redes cabladas e antenas para comunicar com os MNS (ou outros APs) via rádio. Estes *switches* estabelecem a ponte entre os mundos cablado e sem fios. Um conjunto de MNs a usar o mesmo AP forma uma BSS (*Basic Service Set*). Este conjunto pode ser alargado de forma a englobar outros APs, constituindo uma ESS (*Extended Service Set*). A interligação entre vários APs designa-se

por sistema de distribuição (*Distribution System*, DS) e pode ser concretizado por uma rede cablada ou sem fios. Este trabalho concentra toda a sua atenção em redes sem fios 802.11 Infra-Estruturadas.

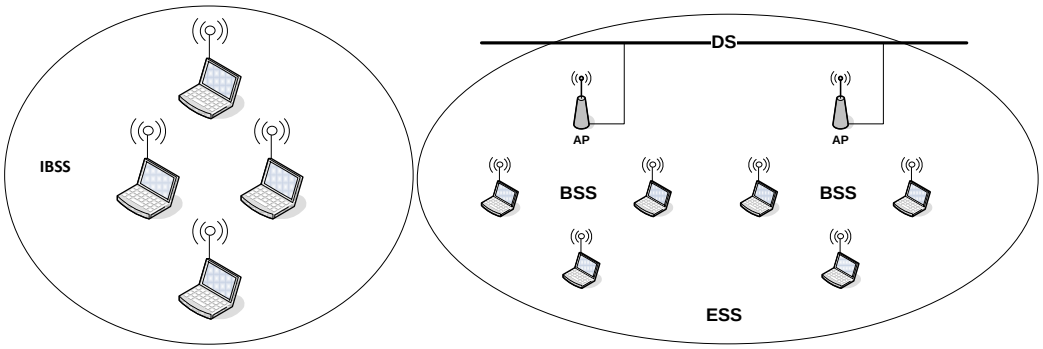

Figura 2.1: Arquitecturas de rede 802.11: Ad-Hoc (IBSS) à esquerda e Infra-Estruturada (BSS e ESS) à direita.

2.1.2 - Comunicação em Redes Infra-Estruturadas

Identificadores de Rede: As redes sem fios identificam-se por identificadores de 32 octetos, o SSID (Service Set ID). Este identificador é estabelecido e imposto ao nível do AP.

Tramas: A comunicação entre um MN e um AP é feita via tramas de vários tipos. Há 3 tipos base de tramas, cada um com diversos subtipos, todos descriminados pelo campo *Frame Control* do seu cabeçalho.

Tipo de Trama	**Nome do Emissor**	
	MN	**AP**
Gestão	-	*Beacon*
	Probe Request	*Probe Response*
	Authentication Request	Authentication Response
	Deauthentication	
	Association Request	*Association Response*
	Reassociation Request	*Reassociation Response*
	Disassociation	
Controlo	Request to Send (RTS)	
	Clear to Send (CTS)	
	Acknowledgment (ACK)	
Dados	Data Frame	

Tabela 2.1: Tipos de tramas e entidade responsável pela sua emissão.

Existem portanto três tipos base de tramas:

• ***Tramas de dados***: Para efectuar troca de dados, geralmente pacotes IP. Cada trama transporta, como dados únicos, o que se designa como MPDU (*MAC Protocol Data Unit*), que pode ser um fragmento de algo maior, um MSDU (MAC Service Data Unit). A informação relativa à fragmentação de um MSDU em vários MPDU, que é relevante para conduzir a desfragmentação, está presente no cabeçalho da trama.

Cabeçalho (30 octetos)							**Dados (MPDU) (0-2312)**	**FCS (4)**
Frame Control (2)	**Duration ID (2)**	**Addr 1 (6)**	**Addr 2 (6)**	**Addr 3 (6)**	**Sequence Control (2)**	**Addr 4 (6)**		

Figura 2.2: Formato de uma trama de dados. Entre parênteses o tamanho em octetos.

• ***Tramas de Gestão***: Permitem a negociação e manutenção da ligação entre o MN e um AP. Os diferentes tipos de tramas de gestão são apresentados de seguida:

***Beacon*:** Anuncia a presença de uma rede e diversas características desta. São enviados periodicamente pelo AP.

***Probe Request*/Response:** Trocadas entre MN e AP quando o MN pede informação sobre um determinado AP. O AP responde com a informação de capacidade, velocidades de transmissão suportadas, etc.

Authentication Request/Response: A autenticação é um processo onde o ponto de acesso aceita ou rejeita a identificação de um MN.

Deauthentication: Trama trocada entre MN e AP para permitir ao MN rescindir uma autenticação anterior ou permitir ao AP terminar uma sessão autenticada com o MN.

***Association Request*:** Permite ao MN pedir a alocação de recursos a um AP para iniciar uma sessão de comunicação.

***Association response*:** Resposta do AP ao pedido de associação do MN. Se o AP aceitar a associação esta trama incluirá informação relativa ao processo, como o ID de associação e velocidades de transmissão suportadas.

***Reassociation Request/Response*:** Se um MN se desloca para um local onde o sinal de outro AP é mais forte, ele usa estas tramas para se associar a este novo AP. Partindo deste ponto o novo AP controla os pacotes de dados.

***Disassociation*:** Serve para expressar a vontade do MN se desassociar de um AP, ou porque se está a desligar ou a mudar de AP. O AP liberta os recursos para serem usados para outra associação de um MN.

• ***Tramas de Controlo***: Servem para gerir a comunicação entre o MN e o AP. Usadas para gerir o acesso ao meio e evitar colisões resultantes de comunicações em simultâneo. Servem também para confirmar a correcta recepção de tramas, devido ao meio de comunicação ser tipicamente ruidoso. A função RTS/CTS é opcional e característica do CSMA/CA (*Carrier Sense Multiple Access with Collision Avoidance*) para reduzir o número de colisões observadas no caso de nós escondidos.

Cabeçalho (30 octetos)						**Dados (MPDU) (0-2312)**	**FCS (4)**
Frame Control (2)	**Duration ID (2)**	**Dst Addr (6)**	**Src Addr (6)**	**BSSID (6)**	**Sequence Control (2)**		

Figura 2.3: Formato de uma trama de controlo. Entre parênteses o tamanho em octetos.

***RTS (Request to Send)*:** Enviado por uma estação na primeira fase de um acordo bidireccional necessário antes do envio de tramas de dados.

***CTS (Clear to Send)*:** Uma estação responde com CTS atribuindo a permissão para o pedido da outra estação no acordo para envio de tramas de dados.

ACK (Acknowledgement): Depois da recepção de uma trama de dados a estação receptora verifica a existência de erros. A estação receptora envia então um ACK para a estação emissora se não houver nenhum erro naquela transmissão. Se a estação emissora não receber nenhum ACK depois de um determinado período retransmite a trama de dados.

2.1.3 - Modo de Conservação de Energia (Power Save Mode) em Redes Infra-Estruturadas

Por omissão, as redes sem fios operam num modo de gestão de energia denominado Modo de Acesso Constante (*Constant Access Mode*, CAM) para ouvir continuamente a rede e obter a informação necessária e oportuna. Contudo, quando o consumo de energia é um problema, como é o caso dos dispositivos móveis, os MNs e APs podem ser configurados com PAM (*Polled Access Mode* ou PS, *Power Save Mode*). Assim, os clientes na rede entram em modo *sleep* (estado de dormência) desligando a NIC (*Network Interface Card*) não consumindo energia.

Em instantes regulares os dispositivos "acordam" para receber um pacote especial chamado TIM (*Traffic Information Map*) que é enviado com cada *beacon* emitido pelo AP. Entre os intervalos de tempo em que os TIMs são emitidos, o cliente desliga a NIC com o propósito de conservar energia, embora mantendo uma sessão activa. Todos os dispositivos na rede partilham o mesmo período de *wake-up* onde acordam para ouvir o TIM pois devem estar activos a tempo de o receber. O relógio

TSF (*Timing Synchronization Function*) corre enquanto os MNs estão no modo *sleep* e garante a sincronização dos MNs.

O TIM informa a existência de informação à espera de ser entregue a cada MN. A NIC do cliente permanece activa quando é recebida no TIM a informação de que existem tramas a serem transmitidas no tampão do AP. Quando esses dados são transmitidos, a NIC volta ao estado de escuta intercalada dos TIMs.

O AP guarda tramas individualmente para cada NIC no buffer até receber uma mensagem de *poll request* proveniente do MN a que se destina a informação armazenada anunciando que o MN respectivo está preparado para receber os dados. O AP anuncia a existência de dados em difusão (broadcast/multicast) através de um pacote DTIM (*Delivery traffic Information Map*). O tempo de emissão do DTIM é sempre um múltiplo do equivalente ao TIM e é ajustável no AP.

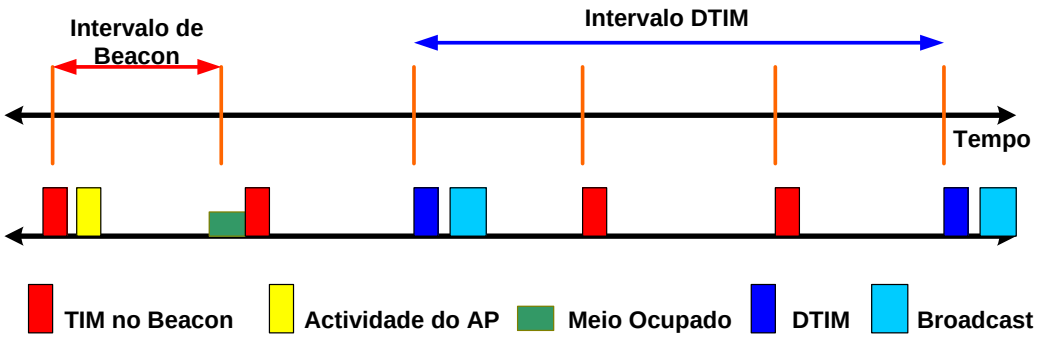

Figura 2.4: Representação temporal da emissão de TIMs, DTIMs, transmissão de dados após a notificação TIM e período de escuta periódico dos TIMs.

A resposta do AP ao *poll request* do MN pode ser imediata e tal é geralmente feito se a trama de dados é grande. Esta resposta pode também ser feita de forma fragmentada, onde dados volumosos são fragmentados em tramas mais curtas. Esta resposta ao *poll request* pode também ser um simples ACK, onde os dados não são enviados imediatamente. O MN deve manter-se acordado até receber os dados até que o seu bit no TIM do *beacon* estar liberto.

2.2 - Serviços da Camada MAC

Os dispositivos que implementam a camada física e MAC IEEE 802.11 como parte de uma WLAN são denominados por estações. Estas estações podem ser MNs ou APs. Os APs são estações que integram o DS e facilitam a distribuição de dados entre MNs. A camada MAC disponibiliza nove serviços lógicos: autenticação, desautenticação, associação, desassociação, reassociação, distribuição, integração, privacidade e entrega de dados. O AP usa todos estes serviços. Cada serviço utiliza um conjunto de mensagens com elementos de informação (*Information Elements*, IEs).

2.2.1 - Os Serviços

• **Autenticação:** Como as WLANs têm segurança limitada na camada física para evitar o acesso não autorizado, o 802.11 define serviços de autenticação para controlar o acesso à rede. A finalidade do serviço de autenticação é proporcionar controlo de acesso à rede cablada LAN. O AP pode exigir a prova de que o MN pertença a um certo utilizador. Sem esta prova de identidade o MN não pode usufruir do serviço de entrega de dados. Todos os MNs devem autenticar-se antes de comunicar com os elementos da rede. A autenticação é feita mediante a troca de tramas de *Authentication Request/Response*, que pode envolver um número arbitrário de tramas numeradas sequencialmente por um número de ordem. Tanto o AP como o MN são configurados pelo administrador com protocolos de autenticação iguais. Na autenticação o AP aceita ou rejeita a identificação do MN consoante a concordância ou não de chaves de cifra. O MN inicia o processo com uma mensagem de *Authentication Request* indicando o modelo de autenticação pretendido (OSA ou SKA) onde o AP responde com um erro caso não seja permitido. Caso seja permitido seguem-se as mensagens de reposta necessárias para conclusão do processo.

Open System Authentication (OSA): Método simples, processado em dois passos, utilizado por omissão. Um MN, com a intenção de se autenticar com outra estação, envia uma trama de autenticação contendo a sua identidade. A estação receptora responde com outra trama notificando a estação que pretende ser autenticada se a sua identidade é reconhecida ou não. Não existe qualquer reforço de segurança, qualquer dispositivo pode usufruir dos serviços conferidos pela rede.

Shared Key Authentication (SKA): Este processo pressupõe uma pré-distribuição de chavesPSK (*Pre-Shared-Key*) ao MN e ao AP. As chaves configuradas no AP podem ser associadas ao SSID ou, mais personalizadamente, associada ao MAC de cada MN autorizado. A recepção prévia da chave é feita através de um canal seguro independente da rede 802.11. O uso de uma chave secreta partilhada requer uma implementação de cifra através do algoritmo WEP (*Wired Equivalent PrivacyI*). A autenticação é então feita por um processo de desafio-resposta, ou seja, o MN envia uma trama de *Authentication Request* ao AP, o AP envia o desafio ao MN numa trama *Authentication Response*, que consiste num conjunto de 128 octetos aleatórios. O MN deverá enviar outra trama *Authentication Request* com uma cópia do desafio mas com a protecção WEP. Assim o desafio é cifrado com uma chave PSK que deverá ser comum ao MN e AP. Após decifrar o recebido do MN, o AP compara com o conjunto enviado, caso o conteúdo seja igual o AP envia uma mensagem A*uthentication Response* com uma autorização de acesso, caso contrário será enviada a mesma trama mas com indicação de uma falha na fase de autenticação.

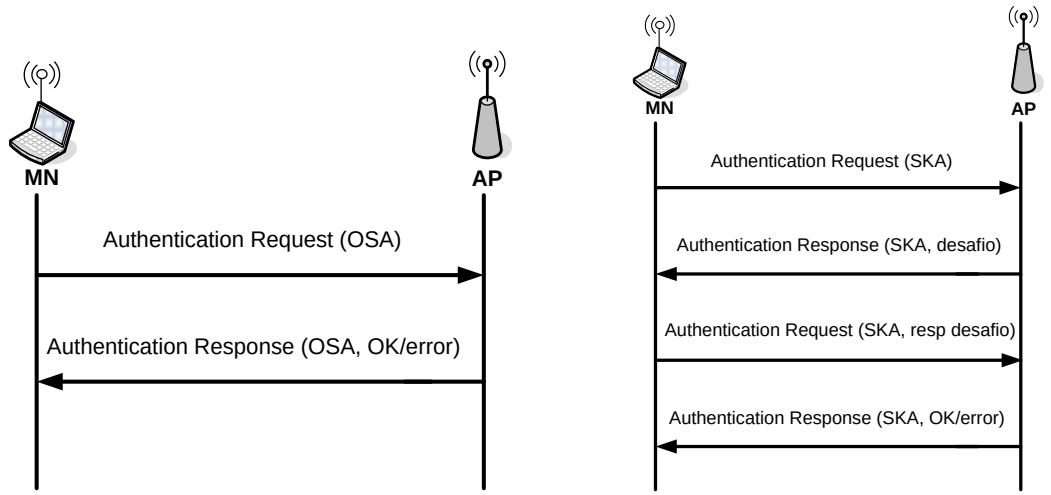

Figura 2.5: Modelos de autenticação WEP: OSA à esquerda e SKA à direita.

• **Desautenticação:** Este processo remove qualquer estado relativo a uma autenticação. Este serviço é usado para eliminar da rede um utilizador previamente autorizado, impedindo-o de fazer uso de qualquer recurso da rede. Para voltar a usar a rede novamente, a estação terá de realizar uma nova autenticação. A desautenticação é uma notificação e não pode ser recusada. Este processo pode ser iniciado tanto pelo MN, no caso de expressar explicitamente a sua intenção de se retirar da rede, ou pelo AP, caso o AP não vá continuar a servir uma rede, ou vai ser desligado ou simplesmente tem alguma razão para desautenticar um MN.

• **Associação:** Após uma etapa bem sucedida de autenticação segue-se a associação, na qual o MN se associa ao AP, o que significa na prática que o AP reserva recursos para identificar e gerir a comunicação com o MN. Com a associação é estabelecida uma ligação lógica entre o dispositivo móvel e o AP. Cada MN deve associar-se a um AP antes de poder enviar dados via AP para o DS. A associação é apenas invocada uma vez, geralmente quando o MN entra numa BSS. O MN pode estar autenticado em vários APs mas apenas pode estar associado a um. Esta autenticação em vários dispositivos permite acelerar a migração de associações de um MN entre vários APs, o que é útil numa situação de movimento do dispositivo. Cada AP pode ter vários MNs a ele associados. A associação é feita por troca de mensagens *Association Request/Response*. Nestas é referido o SSID da rede sem fios a que se pretende ligar, bem como vários parâmetros operacionais.

• **Desassociação:** A desassociação do MN da rede, quando não é feita no âmbito de uma reassociação explícita, pode ser comunicada por qualquer um dos interlocutores através de uma trama *Disassociation*. A desassociação é uma notificação, pelo que nenhum dos intervenientes pode recusar a terminação da associação. Este passo liberta recursos no AP e permite ao MN associar-se a outro AP. Pode ser invocada por um AP para forçar um MN a eliminar a associação consigo ou pode ser enviada pelo MN para informar o AP que, doravante, não irá disponibilizar dos recursos do DS. Quando um MN se encontra desassociado deverá iniciar uma nova associação para comunicar

novamente com um AP. Um AP pode forçar a desassociação de um MN devido a restrições de recursos, porque está a ser desligado ou porque se está a retirar da rede. Um MN deve desassociar-se quando se retira da rede, contudo, nada na arquitectura 802.11 obriga a que tal notificação aconteça.

• **Reassociação**: A associação de um MN pode migrar para outro AP de outra BSS pertencente à mesma ESS a qualquer altura, por iniciativa própria. A reassociação transfere o estado de associação de um MN para outros APs. É feito um pedido de reassociação com troca de mensagens R*eassociation Request/Response* que diferem das tramas de associação por terem um identificador do AP a que o MN estava previamente associado. O MN usa a reassociação repetidamente enquanto se desloca dentro da sua ESS, quando perde o contacto com o AP ao qual está associado e precisa associar-se a outro. O facto de o MN disponibilizar ao novo AP informação sobre o AP anterior permite a comunicação entre esses APs para obtenção de tramas não entregues destinadas ao MN bem como informação relevante à nova associação. É sempre o MN que inicia a reassociação.

• **Privacidade / Confidencialidade:** Numa rede sem fios onde a informação é difundida pelo ar, isto é, não existe salvaguarda da integridade e privacidade no meio físico de transmissão, todos os dispositivos com um receptor rádio podem ouvir as conversações do meio, criando um grave impacto na segurança da informação transmitida. O padrão 802.11 disponibiliza uma solução para este problema oferecendo um serviço de privacidade que aumenta o nível de segurança das transmissões a um estatuto equivalente ao de redes cabladas. O WEP evita a visualização não autorizada das transmissões protegendo os dados enquanto estes circulam no meio sem fios. O serviço de privacidade, aplicado a todas as tramas de dados e algumas tramas de gestão, é um algoritmo de cifra baseado no 802.11b.

> **WEP *(Wireless Equivalent Privacy)***: Para proteger os dados enviados através das WLANs, a norma 802.11b define o uso do protocolo WEP. Este protocolo tenta implementar uma rede cablada numa rede sem fios cifrando os dados nas camadas mais baixas do modelo OSI. O protocolo WEP está baseado no algoritmo de cifra RC4 e utiliza chaves de 64 bits ou de 128 bits. Na verdade são de 40 a 104 bits, já que os outros 24 bits vão no pacote como Vector de Inicialização (*Initialization Vector*, IV). Este protocolo permite a autenticação unidireccional dos MNs, a confidencialidade e o controlo de integridade dos dados trocados entre o MN e o AP. O WEP inclui duas funcionalidades: autenticação de MN e confidencialidade e controlo de integridade dos dados trocados. A integridade e confidencialidade são implementadas por meio de cifra de conteúdos com a chave pré-partilhada. A cifra é apenas aplicada a tramas unicast, não existindo nenhum suporte para protecção de tramas em difusão (*multicast* ou *broadcast)*. O nível de segurança implementado mediante este protocolo é reduzido. Um atacante que capture o desafio e a resposta pode calcular a chave e autenticar-se a si mesmo.

• **Distribuição:** A distribuição é o serviço principal usado por estações 802.11, e fá-lo a cada vez que envia uma trama MAC através do DS. Os três serviços de associação (associação, reassociação, desassociação) conferem a informação necessária à operação do serviço de distribuição.

• **Entrega de Dados:** Permite a transferência de dados entre as estações num ambiente sem fios.

• **Integração:** O serviço de integração liga a WLAN a outras LANs, incluindo uma ou mais LANs cabladas ou WLANs 802.11. Permite a transferência de dados entre o DS de uma LAN 802.11 e uma LAN que pode ou não ser 802.11. Faz a tradução das tramas 802.11 para tramas que podem atravessar outras redes. A estação que disponibiliza esta funcionalidade é denominada de *portal*. O portal é um conceito de estrutura abstracta que tipicamente reside no AP, apesar de poder fazer inteiramente parte de outra estrutura de rede.

2.2.2 - Máquina de estados do MN

Quando um MN se quer ligar a uma WLAN, sabendo o seu SSID, completa este processo em dois passos: autenticação e associação.

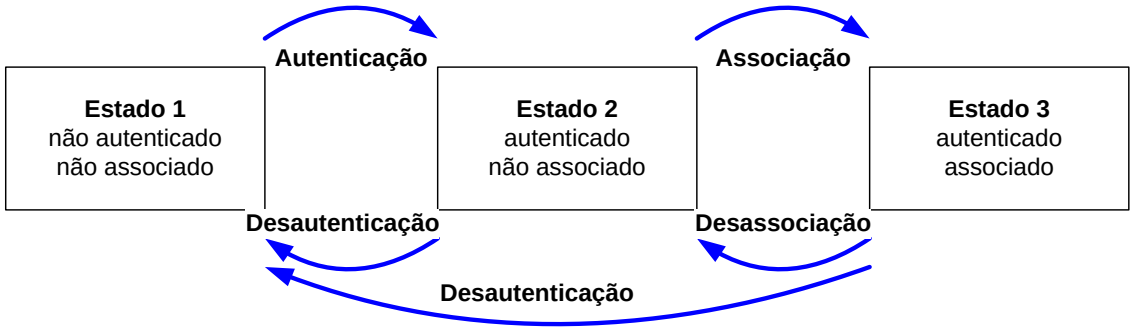

Figura 2.6: Máquina de estados de um MN 802.11.

Estes estados são usados para determinar que tipo de tramas podem ser trocadas entre o MN e o AP a cada momento do processo de associação. Para passar do estado 1 ao 3 deve se passar pelo estado 2, isto é, a autenticação deve sempre preceder a associação. O MN pode manter o estado 2 em vários APs mas apenas pode efectuar a transição deste estado para o estado 3 com um só AP de cada vez. Do estado 3 pode passar directamente ao estado 1, não autenticado nem associado, através de uma trama *Deauthentication*.

2.3 - *Handover*, *Handoff*, Roaming e Mobilidade

Estes termos significam essencialmente o mesmo. Contudo, mobilidade é o termo mais utilizado em redes cabladas, particularmente em redes IP, enquanto os outros termos são aplicados em sistemas sem fios. Neste trabalho, *handover*, *handoff*, roaming e mobilidade serão usados sempre com o mesmo significado: transição de AP por parte de um MN. Conseguir o *handoff* intra-domínio e inter-domínio é um desafio que deve ser levado em atenção no desenvolvimento actual das WLANs.

A facilidade de uso e o baixo custo de redes sem fios infra-estruturadas tornou-as a forma mais popular de acesso à internet. Esta popularidade implicou uma densificação do uso destas infra-estruturas proporcionando um conjunto de possibilidades de acesso pelo qual o utilizador pode optar. Esta densificação proporciona uma conectividade contínua mesmo a utilizadores altamente móveis. A distância de alcance do AP é limitada e leva a *handoff*s frequentes, bem como a necessidade de melhor qualidade de sinal ou menor carga de rede de cada AP.

Para suportar utilizadores móveis dentro de uma rede 802.11 são necessárias técnicas eficientes para *handoff* transparente entre APs. É necessário manter a ilusão de conectividade contínua sendo necessário escolher o momento correcto para realizar o *handoff* e escolher o AP correcto para o qual mudar. Um *handoff* demorado resulta na perda de pacotes, atrasos, jitter (variação do atraso), retransmissões e outros aspectos indesejáveis. Serviços em tempo real requerem tempos de *handoff* que estão ainda aquém do ideal para os suportar.

2.3.1 - Suporte de mobilidade

A libertação do meio físico em cobre (cabos) e a mobilidade são as principais motivações da implementação de redes sem fios. As estações podem deslocar-se livremente mantendo o estado ligado com a rede e transmitindo tramas durante o movimento. A mobilidade pode originar três tipos de transição:

• **Sem Transição:** Quando o movimento do MN não se propaga além da área de cobertura do AP, não se verifica transição. Apesar de isto não se tratar de uma transição, é assim definida em [4].

• **Transição de BSS:** este tipo de transição é definido como o movimento de um MN da área de cobertura de um AP para a área de cobertura de outro AP, havendo transferência de associação de um para o outro. Ambos os APs, integrantes de duas BSSs distintas, pertencem à mesma ESS.

• **Transição de ESS:** Este tipo de transição reflecte-se no movimento de um MN partindo de uma BSS enquadrada numa ESS para a BSS de outra ESS. O 802.11 não suporta este tipo de transições, excepto para permitir a associação do MN no AP da ESS destino. O mais provável de ocorrer é a interrupção das ligações de nível superior. Para manter as ligações de nível superior é preciso o suporte dos protocolos em questão. No caso do TCP/IP, é necessário utilizar o Mobile IP [5] para permite uma

transição de ESS transparente. Geralmente uma rede WLAN encontra-se dentro de uma ESS e numa sub-rede IP. Isto poderá implicar o seguinte:

a) *Handoff* Inter-Subrede.

b) *Handoff* Inter-Domínio, como por exemplo entre duas redes diferentes.

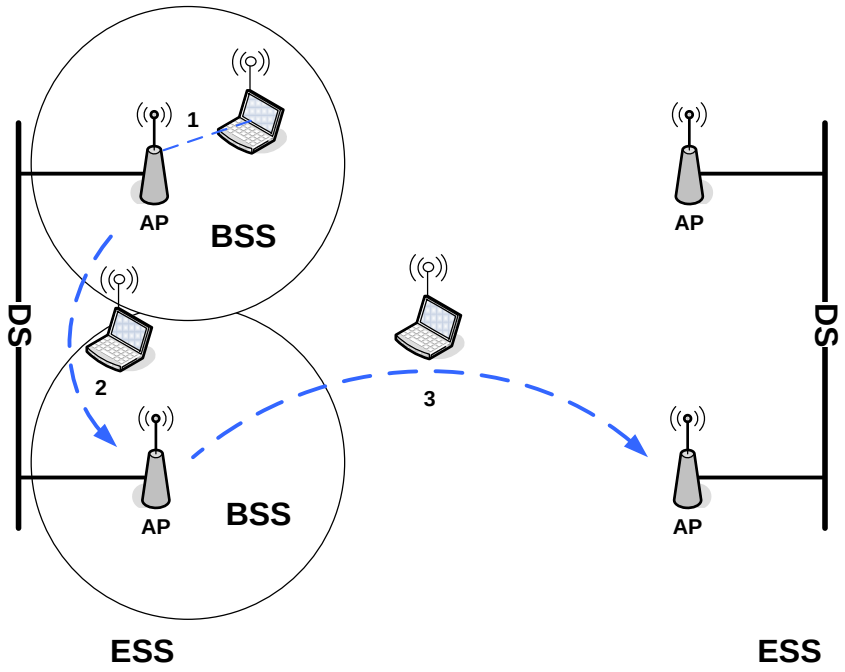

Figura 2.7: Diferentes tipos de transição: **1**- Sem Transição; **2-** Transição de BSS; **3-** Transição de ESS.

2.3.2 - Noções de *Handoff*

O principal benefício de redes sem fios consiste no facto de um cliente permanecer ligado enquanto se movimenta. Para manter a ilusão de conectividade contínua para o cliente é implementado um processo chave, o *handoff*, que consiste num conjunto de técnicas envolvidas para direccionar o tráfego destinado a um determinado utilizador de um emissor para outro. Idealmente o *handoff* é transparente ao cliente, sem que este se aperceba que tenha ocorrido. Contudo esta técnica ainda não foi dominada por completo.

Utilizadores com dispositivos móveis apercebem-se sempre que ocorre um *handoff* devido a interrupções e perda temporária de conectividade. Portanto o desafio permanece, proporcionar mobilidade contínua sem interrupção de comunicação, isto é, *handoff* transparente (como

transparente entende-se que passa absolutamente despercebido à percepção do utilizador do dispositivo móvel).

Podem distinguir-se duas modalidades do processo de *handoff*:

• ***Hard Handoff***: também conhecido por *break-before-make*, implica uma quebra abrupta da conectividade durante o tempo que demora o reencaminhamento do tráfego de um utilizador para uma diferente infra-estrutura de rede.

• ***Soft Handoff***: também denominado por *make-before-break*. Não interrompe a ligação enquanto não for criada a nova associação que passará a ser usada pelo utilizador. Em contraste com o *hard handoff*, esta modalidade do processo implica que ambos os APs possam simultaneamente receber o tráfego do dispositivo móvel. Geralmente o *soft handoff* irá proporcionar uma probabilidade superior de implementar um *handoff* transparente, contudo não é facilmente implementável, tanto devido à tecnologia específica em uso como também devido ao local onde o *handoff* ocorre na topologia da rede. A tecnologia 802.11 sempre foi construída assente no conceito de *hard handoff*.

2.3.3 - Fases do *handoff* 802.11

Variando consoante o autor, o *handoff* pode decompor-se em quatro fases:

• **Descoberta:** Não é definido o instante em que esta fase tem lugar, podendo ser efectuada intercalada com a comunicação normal, ou integrante do processo de *handoff* levado a cabo por completo no momento em que este é accionado. Caso o MN não tenha outra forma de obter informação sobre os APs a servir a rede nos diversos canais (ex: *Neighbour Graphs*, entidade responsável por difundir essa informação, etc.), esta é a forma que terá de ser utilizada para analisar o meio onde o MN se encontra inserido.

O cliente da rede faz a recolha de informação necessária analisando o meio onde se encontra inserido, sintoniza a sua interface de rádio em todos os canais, um por um, e obtém informação dos APs ao alcance bem como as suas métricas (elementos de informação). Quando um MN pretende ligar-se a uma rede sem fios deve primeiro saber qual o SSID desta.

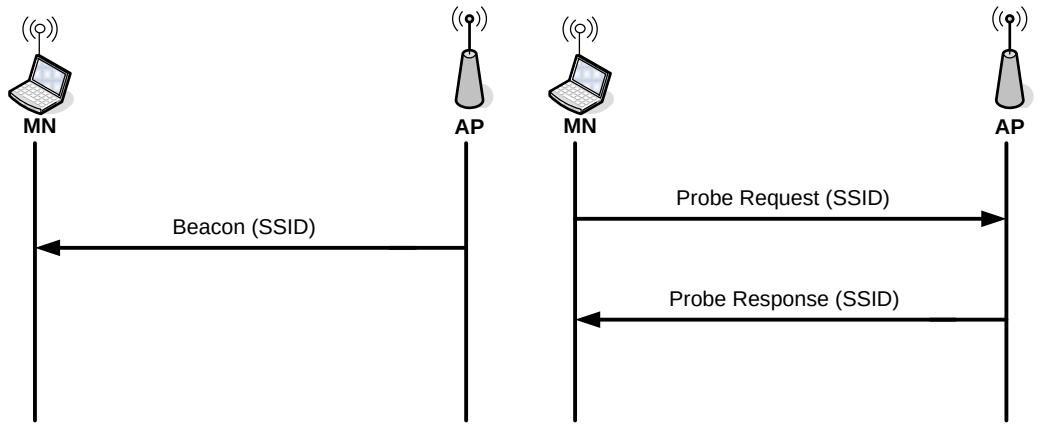

Figura 2.8: Tramas trocadas durante: à esquerda aquisição passiva, à direita aquisição activa.

Existem duas formas distintas de o MN realizar a perscrutação do meio:

> **Aquisição Passiva (*Passive Scanning*):** Consiste em perscrutar passivamente todos os canais de comunicação e receber tramas de gestão *beacon*. Estas são periodicamente enviadas pelo AP para anunciarem, por entre outras coisas, os SSID das redes a que pertencem. O MN regista a potência do sinal calculada sobre esse *beacon*.
>
> **Aquisição Activa (*Active Scanning*):** Interrogação do AP, com uma trama *Probe Request*, para saber se serve ou não uma rede com um determinado SSID. É enunciado claramente nas tramas de *Probe Request* qual o SSID que está a ser sondado. Esta sondagem é feita em todos os canais de comunicação. A NIC envia uma trama *Probe Request* e todos os APs no canal respondem com um *Probe Response*. O MN espera durante um período denominado *MinChannelTime* pela resposta de algum AP no canal actualmente a ser sondado. Caso nenhum AP responda durante este período o MN sonda o canal seguinte. Caso seja recebida uma resposta no canal, o MN espera então um período *MaxChannelTime* por outros APs nesse canal que possam ainda responder.

• **Accionamento do Processo de *Handoff*:** Nesta fase o MN toma consciência da necessidade de se associar a outro AP e acciona o processo propriamente dito. Consiste na altura temporal em que um MN identifica a necessidade de procurar outras associações a APs. A implementação de mecanismos de decisão apropriados é deixada ao critério do fabricante das NIC (soluções proprietárias implementadas no firmware), mas genericamente, segue um modelo de accionamento do processo de handoff no instante em que a intensidade de sinal do AP actual medido pelo MN decai abaixo de um offset específico que despoleta a perscrutação. Actualmente esta decisão leva em conta os seguintes critérios:

- Não recepção de um número específico de tramas de ACK;
- Perda de um certo número de *beacon*s consecutivos.

• Perda de tramas de *beacon*.

• Redução da potência de sinal abaixo de um limiar específico.

Neste último caso, o driver da NIC acciona o processo quando a qualidade de sinal decai abaixo de um limiar definido. Se a fase de descoberta tiver de ser efectuada durante o *handoff*, é então definido um limiar inicial, chamado *cell search threshold*, que difere do limiar de *handoff* por ter um valor mais elevado. A transição para o novo AP só ocorre, porém, quando há decaimento abaixo do *limiar de handoff*. Velayos e Karlsson [6] apresentam evidências de que uma decisão mais simples, baseada na perda consecutiva de 3 pacotes, é mais rápida e eficaz. A figura seguinte exemplifica o conceito de limiar aplicado no *handoff* 802.11.

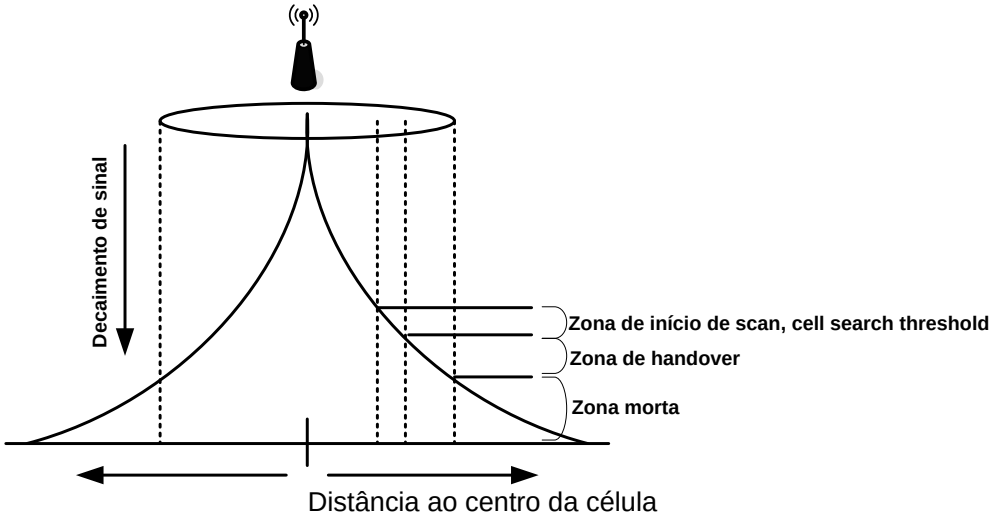

Figura 2.9: Zonas onde são disparados os processos de pesquisa da rede e *handoff* na relação intensidade de sinal em função da distância geográfica ao AP em linha recta.

• **Selecção do AP:** Quer a informação sobre o meio tenha sido obtida pelo processo de descoberta de forma proactiva ou durante o processo de *handoff*, quer tenha sido fornecida por qualquer outra entidade, é necessário escolher, de entre as opções válidas existentes no meio circundante, qual o AP mais apropriado para destino do *handoff* eminente. Actualmente essa selecção é feita com base no RSSI (*Received Signal Strength Indicator*) no cliente, ou seja, num universo de APs, o escolhido será o que apresenta melhor RSSI. Pode também ser escolhido o AP mais apropriado tendo em conta o SNR (*Signal to Noise Ratio*) recebido nas tramas *Beacon* e *Probe Response*. Um AP apenas será favorável se o SNR medido for superior, pelo menos Δ (histérese), que o SNR medido no AP actual. Esta histerese é usada para evitar operações de *handoff* desnecessárias que poderão desencadear um efeito de "ping--pong" quando um cliente é servido com eficácia equivalente por diferentes APs. Este Δ é definido pelo fabricante e depende de cada solução proprietária.

• **Compromisso:** Este é o *handoff* em si. Nesta fase final do *handoff* o cliente desassocia-se do AP actual que o serve e formaliza o acordo de serviço com o novo AP previamente escolhido como sendo o destino mais apropriado. É efectuada a autenticação, que é o processo segundo o qual o cliente comprova a sua identidade ao AP. Segue-se a associação onde são trocadas tramas entre o AP e o cliente resultando na reserva de recursos por parte do AP e preparação do acesso do cliente à rede. Por fim a rede cablada deverá ser informada sobre o *handoff* e instruída para que os pacotes sejam redireccionados para o novo AP. Geralmente esta parte é levada a cabo pelo protocolo de *spanning tree* 802.1d. No caso de estar definida segurança através da aplicação do padrão 802.11i [7] a autenticação 802.11 será Open System, e posteriormente é feita a autenticação mútua entre o MN e o novo AP usando 802.1X [8]. Se o *handoff* for efectuado entre diferentes sub-redes ou entre domínios diferentes, deve ainda considerar-se o atraso do *handoff* da camada 3, nomeadamente o tempo de execução dos protocolos DHCP [9] ou MIP [5].

2.4 - IEEE 802.11F: Práticas recomendadas para IAPP (Inter Access Point Protocol)

A recomendação IEEE 802.11F [10] descreve uma extensão opcional ao IEEE 802.11 que proporciona comunicação entre APs de diferentes fabricantes. Os objectivos principais do IAPP são os seguintes:

- Manter uma única associação invariável numa rede sem fios.

- Transferência segura de estado e contexto entre APs envolvidos numa reassociação. Este contexto pode conter informação relativa a fluxos IP, contexto de segurança, QoS, compressões de cabeçalhos e informação AAA (*Authentication, Authorization and Accounting*).

2.4.1 - Noções Sobre o seu Funcionamento

O IEEE 802.11F define uma trama de *Update* para a camada protocolar 2. Esta trama é construída para ser difundida pelo novo AP (alvo da transição originada pelo processo de *handoff*) a todos os dispositivos da camada 2 dentro da mesma sub-rede. O MN é indicado como fonte nesta trama. O objectivo do seu envio é actualizar a tabela de encaminhamento nas *bridges* MAC e *switches* da camada 2 existente na rede para que futuros pacotes endereçados ao MN sejam encaminhados pelo novo AP actualmente a servi-lo.

Tal como definido no IEEE 802.11F, quando um MN se reassocia, o novo AP envia uma trama *Update* de camada 2 antes de qualquer outra coisa. Como esta trama é difundida em broadcast, o AP anterior irá receber a trama. O *driver* da NIC sem fios do AP anterior pode deduzir que o MN esteve antes na

sua BSS. A recepção desta trama informa portanto o AP antigo que o MN se deslocou para outro AP. Isto garante uma única associação para cada cliente na rede.

No início da reassociação, o novo AP pode opcionalmente enviar uma mensagem *Security Block* ao AP anterior, que confirma a recepção desta mensagem enviando um *Ack-Security-Block*. Esta mensagem inclui informação de segurança para estabelecimento de um canal com comunicação seguro entre os APs. O novo AP envia um *Move-Notify* ao AP anterior pedindo informação de contexto sobre o cliente e notificando-o sobre a reassociação. O AP anterior responde com um *Move-Response*.

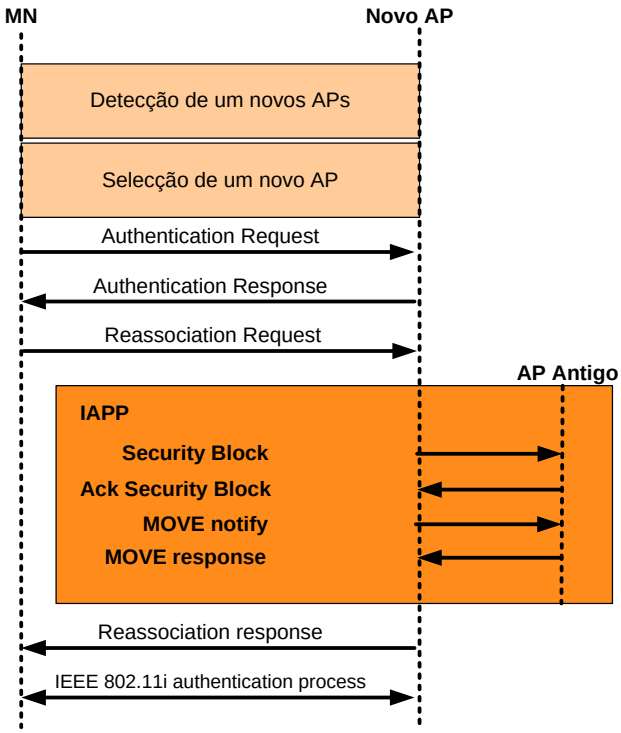

Figura 2.10: Reassociação com uso de IAPP

Para garantir confidencialidade da informação de contexto trocada, o IAPP recomenda o uso de um servidor RADIUS (para obtenção de chaves partilhadas) para assegurar segurança na comunicação entre APs. O servidor RADIUS pode também disponibilizar o mapeamento de endereços entre endereços MAC e endereços IP, necessários para comunicação entre APs com IAPP na camada de rede.

O IAPP define um mecanismo de *cache* que permite que os APs troquem informação de contexto proactivamente (antes da reassociação). A gestão da *cache* é baseada num *Neighbour Graph* (NG) mantido por cada AP. Após uma associação de um cliente ao AP, este pode transmitir informação sobre esta associação aos APs vizinhos usando uma mensagem *CACHE-notify*. Cada vizinho notificado responde com uma mensagem *CACHE-response* de forma a confirmar que a sua cache foi actualizada.

2.5 - Neighbour Graphs (NG)

2.5.1 - Definição de NG

Um NG é um grafo não orientado onde cada aresta representa o caminho de mobilidade entre APs e os vértices representam os APs numa rede. As arestas interligam APs que têm sobreposição de alcance entre si. A figura seguinte demonstra um exemplo de uma distribuição espacial aleatória de APs e o NG correspondente:

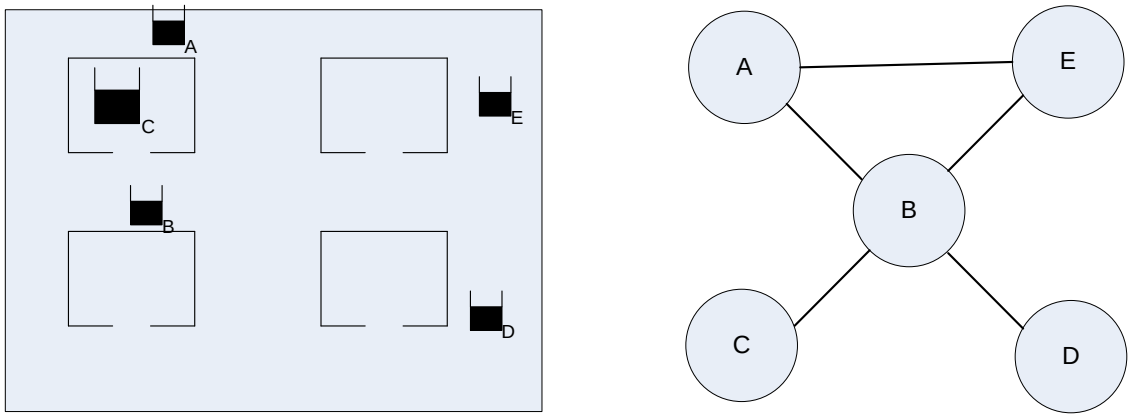

Figura 2.11: Representação espacial de APs de uma rede sem fios à esquerda e respectivo NG à direita.

O grafo não direccionado que representa o NG é definido como:

G = (V , E),
V = (ap1, ap2, …, api),
e = (api, apj),

$$N(api) = \{apik : apik \in V, (api, apik) \in E\}$$

onde G é a estrutura de dados do NG, V é o conjunto de todos os APs, E é o conjunto de arestas e N é o conjunto de APs vizinhos (com ligação não direccionada) de um determinado AP.

Estes grafos reflectem a relação de vizinhança entre APs numa rede, ou seja, o paralelismo entre os APs adjacentes cuja área de cobertura se sobrepõe, permitindo que um MN seja servido por um ou outro AP numa determinada área de sobreposição do serviço de rede. Contêm informação sobre o conjunto de canais onde os APs estão a operar e sobre o conjunto de APs vizinhos em cada um desses canais. Com esta informação o cliente evita sondar canais desnecessariamente e passar menos tempo à espera de respostas de APs inexistentes, ou seja, diminuir o tempo que o MN gasta na fase de descoberta.

2.5.2 - Construção de um NG

O NG pode ser automaticamente gerado de duas formas. A primeira usa mensagens de *Reassociation Request* de um cliente que contém o MAC do AP anterior. Este método decorre da seguinte forma: se um cliente envia um Reassociation Request ao APi, com o MAC do APj antigo, é criada uma relação de vizinhança (i,j) entre estes APs que se reflecte numa aresta no NG. Esta forma de geração de NGs não permite a utilização dos mesmos nos primeiros tempos de vida da rede. A outra forma consiste na actualização de informação entre APs por meio do IAPP (Inter Access Point Protocol).

No contexto de NGs, também podem ser referidos os grafos de canais não sobrepostos (non-overlaping channels) que é uma estrutura de dados que contém informação de relações de não-sobreposição entre APs. Dois APs são não-sobrepostos se e só se um cliente não consegue comunicar com ambos os APs com qualidade de ligação aceitável. Quando dois APs são não-sobrepostos, um *Probe Response* de um indica a impossibilidade de chegar ao outro. Desta forma, durante a fase de descoberta, o MN pode descartar APs ou canais onde os APs estão inatingíveis. Ao receber uma trama de *Probe Response* de APs que se relacionam num grafo de canais não-sobrepostos com outros APs, implica que esses outros APs estão garantidamente fora do alcance do MN. No grafo da figura 2.10, um exemplo de uma entrada do drafo de não-sobreposição seria, por exemplo, a relação AP_A com o AP_D pois um MN que receba resposta do AP_A a um *Probe Request* estaria numa localização geográfica onde nunca iria receber uma resposta do AP_D.

2.6 - IEEE 802.11i

O grupo de trabalho 802.11i empenha-se na implementação de segurança no acesso a redes 802.11. Este esforço foi iniciado como resultado da insatisfação dos utilizadores na exploração de WEP em redes 802.11, que, por volta de 2001, quando o grupo foi criado, era o único método de protecção de ligações a redes 802.11.

Foram, de facto, as extensões de segurança desenvolvidas pelo grupo de trabalho 802.11i [7] que proporcionaram a globalização das redes 802.11. Os termos *Robust Secure Networks* (RSN) e *Safe Secure Networks* (SSN) foram ambos utilizados na apresentação do 802.11i para descrever o seu objectivo principal, contudo o padrão ratificado usa apenas RSN. Os procedimentos RSN são definidos e ocorrem durante a fase de associação, isto permite ao cliente e ao AP determinar o contexto de segurança da sua associação. O contexto de segurança, no seu nível mais básico, irá estabelecer se vai ser usado o modo de segurança *personal* ou *enterprise*. O modo *personal* é chamado *Pre-Shared Key* (PSK) e pressupõe um uso facilitado mas adequado a ambientes SOHO (*Small Office, Home Office*). O modo de segurança *enterprise* usa o padrão 802.1X [8] para autenticação e distribuição de chaves. Este modo é mais robusto que o modo *personal* mas requer um conhecimento mais profundo sobre tecnologias de segurança de forma a permitir o desenvolvimento e implementação do protocolo de

segurança. Também é necessária uma infra-estrutura de autenticação mais complexa. Esta complexidade é a razão do uso do termo "*enterprise*".

A extensão 802.11i apresenta um processo completo de mecanismos de autenticação mútua para o cliente e AP baseado em EAP [11] e 802.1X [8]. Associa este mecanismo a um algoritmo de troca de chaves que permite aos intervenientes o uso de material de cifra dinâmico.

2.6.1 - Segurança de comunicação 802.11i

Independentemente do modo e cifra usados, a segurança 802.11i precisa de duas chaves para proteger as interacções entre MNs e APs: uma PTK (*Pairwise Transient Key*) e uma GTK (*Group Transient Key*). A PTK é usada para proteger o tráfego *unicast* trocado entre o AP e o MN. A GTK é usada na cifra de tráfego em difusão (*broadcast*/*multicast*) enviado a todos os MNs na BSS. A operação básica segue os seguintes passos:

1. O MN associa-se em modo OSA com o AP a servir a rede e são negociados os parâmetros usados com a associação.
2. O AP autentica o utilizador em modo de segurança *enterprise*. Este passo não existe no modo *personal*.
3. Há uma validação e distribuição de chaves num acordo em quatro passos (*Four-Way Handshake Protocol*, 4WHP) para que a PTK se torne disponível no MN e no AP.
4. As chaves temporárias são instaladas, usando a cifra negociada, e as tramas seguintes são protegidas.

Na versão final do 802.11i, correspondente ao WPA2, a GTK cifrada é enviada durante o 4WHP. Na versão precedente, correspondente ao WPA, a GTK era comunicada num acordo em dois passos dedicado a esta instalação da GTK. Após instalada a chave GTK em todos os MNs da BSS, a sua actualização é feita com o acordo em dois passos, quer no WPA como no WPA2.

O passo 3 é equivalente em ambos os modos *personal* e *enterprise*. Nesse passo, o 4WHP é usado para derivar a PTK de uma PMK (*Pairwise Master Key*). A diferença entre os modos reside na fonte da chave PMK. No modo *personal*, a chave PMK é calculada a partir da PSK, que já se encontra presente tanto no AP como no MN antes do passo 3 (PSK é um valor codificado a partir de uma senha conhecida pelo AP e por todos os MNs que esperam associar-se à rede servida por esse AP). No caso do modo de segurança *enterprise*, a PMK é dinamicamente derivada através do processo que ocorre no passo 2. Este método confere um elevado nível de entropia ao modo *enterprise* pois a PMK é renovada a cada sessão do MN.

A figura seguinte apresenta a hierarquia de chaves 802.1X.

MSK :	Master Session Key	**PTK:**	Pairwise Temporary Key
PSK:	Pré-Shared Key	**KCK:**	Key confirmation Key
PMK:	Pairwise Master Key	**KEK:**	Key Encryption Key
PRF-X:	Pseudo-random function producing X bits	**TK:**	Temporary Key

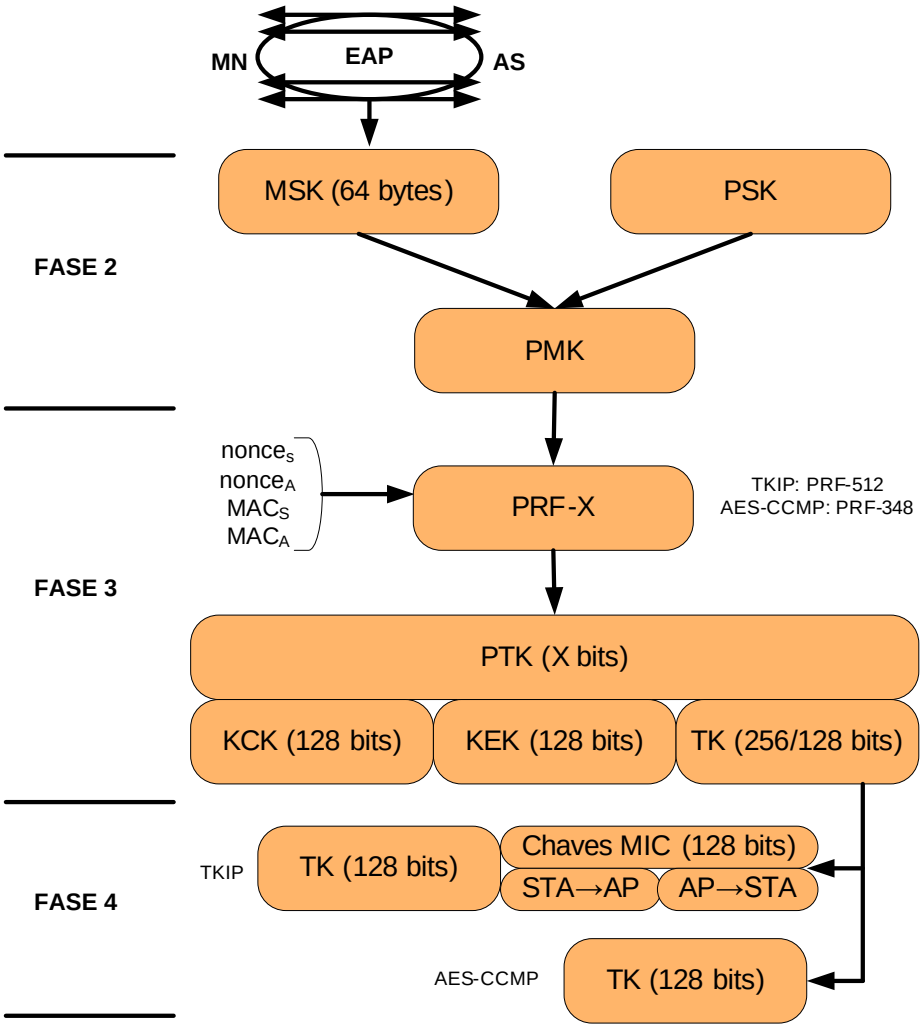

Figura 2.12: Hierarquia de chaves 802.1X.

2.7 - IEEE 802.1X

O 802.1X [8] oferece um protocolo que permite autenticar e autorizar dispositivos ligados à rede. Este proíbe o acesso dos dispositivos à rede enquanto não conseguem autenticar-se com um servidor de autenticação (*Authentication Server*, AS).

Este padrão assegura a segurança através da implementação de mecanismos que se baseiam no conceito de porto controlado/não controlado e controlo de acesso ao nível da camada 2. Os três principais elementos definidos são os seguintes: Suplicante (o MN no caso de redes 802.11), Autenticador (AP no caso de redes 802.11) e Servidor de Autenticação (servidor RADIUS ou qualquer outro servidor AAA). O padrão 802.11i define como usar o 802.1X no contexto de redes sem fios 802.11.

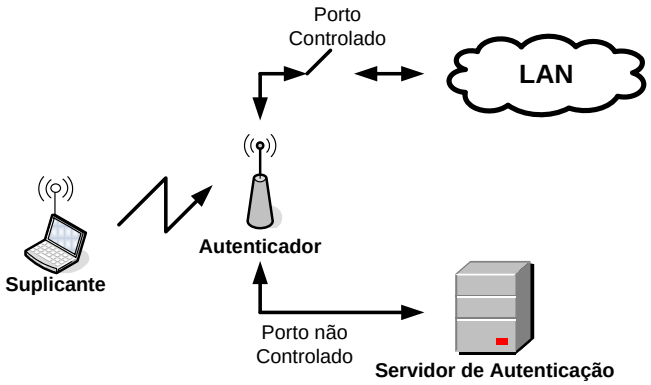

Figura 2.13: Esquema de uma rede com acesso controlado com 802.1X.

O porto não controlado é usado para encaminhar tráfego de autenticação entre o Suplicante e o Servidor de Autenticação. Após uma autenticação bem sucedida, o Servidor de Autenticação informa o Autenticador (AP) sobre o estado da autenticação. De seguida envia todo o material relativo a chaves de autenticação ao Autenticador através de uma troca de chaves com EAPOL. A esta altura o Suplicante e o Autenticador partilham as mesmas chaves, nomeadamente a PMK.

2.7.1 - EAP, Extensible Authentication Protocol

O EAP é um framework de autenticação usado frequentemente em redes sem fios e ligações ponto a ponto. É definido no RFC 3748 e define formatos de mensagens, cada protocolo que use EAP estabelece uma forma de encapsulamento para as mensagens EAP nas suas mensagens protocolares. Os padrões WPA e WPA2 adoptaram oficialmente vários métodos de EAP como mecanismos oficiais de autenticação tais como EAP-TLS, EAP-TTLS, PEAP, LEAP, de entre outros. Quando é invocado o EAP por um servidor de rede com suporte para 802.1X tal como APs 802.11a/b/g, os métodos EAP

avançados podem conferir um mecanismo de autenticação seguro e negociar uma PMK entre o cliente e um servidor de autenticação. Na autenticação MN/AS do 802.11i *Enterprise* permite a criação da chave MSK (Master Session Key) bem como a chave EMSK (*Extended Master Session Key*) que aparece como uma extensão da MSK, mas não é usada no 802.1X.

2.7.2 - O Servidor de Autenticação (AS)

O papel do AS é tomar decisões de autenticação e autorização pela rede. O AS processa as credenciais transmitidas do cliente que requer admissão na rede (Suplicante) e aceita ou rejeita baseado nessas credenciais de acesso a ser pedido. Quando o acesso é concedido, essa autorização pode ser acompanhada pela autorização de diferentes tipos de acesso. O AS pode também iniciar ou parar o registo de informação de contabilização relacionada com a sessão desse utilizador.

2.7.3 - Autenticação do Utilizador com 802.1X em Redes 802.11

O processo de autenticação completo pode ser dividido em três passos principais.

1 - Descoberta e associação 802.11

Nesta fase o Suplicante (MN) liga-se à rede. O MN pode iniciar uma fase normal de descoberta, seguida de autenticação e associação com o AP escolhido. Como esta primeira autenticação não é relevante já que está a ser implementado o 802.1X, usa-se usualmente o OSA para autenticação. No final desta fase o MN está autenticado e associado com o AP mas o porto 802.1X de acesso à rede está fechado (porto controlado fechado).

2 - A autenticação EAP

1. Após detectar a associação 802.11, tanto o Suplicante como o Autenticador podem enviar uma mensagem *EAPOL Start*.
2. O Autenticador abre o porto não controlado para a sessão de autenticação 802.1X, deixando todo o tráfego não 802.1X bloqueado no porto controlado.
3. O Autenticador envia uma mensagem *EAP Request/Identity*.
4. A mensagem *EAP Response* do Suplicante com a identidade do utilizador é passada ao AS na primeira mensagem *RADIUS Access Request*.
5. O AS desafia o Suplicante a comprovar a sua identidade e o AS pode enviar as suas credenciais para comprovar a sua própria identidade ao Suplicante (se o Suplicante exigir autenticação mútua). Esta informação é codificada numa mensagem EAP e enviada ao AP

na carga de uma mensagem *RADIUS Access Challenge*, que é entregue ao Suplicante pelo AP numa trama EAPOL.

6. O Suplicante envia as suas credenciais ao servidor de forma a permitir ao AS verificar a identidade do cliente. As credenciais do cliente estão incluídas numa mensagem EAP e transportadas ao AP numa mensagem EAPOL e dirigidas ao AS pelo AP numa segunda mensagem *Access Request*.

7. No final desta fase o MN partilha uma chave MSK com o AP e o porto 802.1X de acesso à rede continua fechado.

3 - *Four-Way Handshake* Protocol(4WHP)

O envio da chave MSK ao AP fá-lo iniciar o 4WHP com o Suplicante. O processo de autenticação requer o cumprimento de dois processos: o Autenticador e o Suplicante devem autenticar-se mutuamente, e as chaves de cifra do tráfego devem ser derivadas. A troca EAP anterior forneceu a MSK que irá durar a sessão completa e que, por isso, deve ser minimamente exposta. Assim, o 4WHP é usado para estabelecer a PTK, gerada partindo da MSK (depois de transformada em PMK), um *nonce* do AP e outro do MN, e dos endereços MAC do AP e MN. O acordo também estabelece a GTK para cifra das tramas em difusão. As tramas trocadas no 4WHP são apresentadas na Figura 2.14 que representa todo o protocolo.

Tal como demonstrado na Figura 2.12, o ponto de base da hierarquia de chaves é a Master Session *Key* (MSK) gerada separadamente pelo Suplicante e pelo AS. Da MSK é obtida a Pairwise Master Key (PMK) e uma função pseudo-aleatória, PRF-X, é aplicada a esta PMK e outros parâmetros para criar a *Pairwise Transient Key* (PTK). Esta é dividida em 3 chaves. A primeira é a *EAPOL Key Confirmation Key* (KCK), usada nas tramas EAPOL-key para conferir autenticidade. A segunda chave é a *EAPOL Key Encryption Key* (KEK). A KEK é usada na troca EAPOL-key para conferir confidencialidade. A terceira chave é a *Temporary Key* (TK) que é usada pelos protocolos de confidencialidade de dados (TKIP ou CCMP).

O ponto de partida da hierarquia de chaves de grupo é a *Group Master Key* (GMK), que não passa de um valor aleatório. Uma função pseudo-aleatória, PRF-X, é aplicada à GMK e outros parâmetros para criar a *Group Temporal Key* (GTK) usada para cifrar tráfego em difusão.

O maior dilema deste protocolo reside no facto de a autenticação 802.1X ser feita depois da autenticação/associação 802.11, isto significa que o seu atraso associado irá fazer parte da latência de *Handoff*.

Cada uma destas fases é demorada. Experimentalmente foi demonstrado que a autenticação 802.1x demora cerca de 800ms e o 4WH demora cerca de 40ms [12].

A título de resumo apresenta-se a tabela seguinte que retrata resumidamente os métodos de segurança 802.11 existentes e funcionalidades associadas.

<table>
<tr><th colspan="2">Tipo de rede</th><th>Pré-RSN</th><th colspan="2">RSN</th></tr>
<tr><th colspan="2">Funcionalidade</th><th>WEP</th><th>WPA</th><th>802.11i (WPA2)</th></tr>
<tr><th colspan="2">Autenticação</th><td>Unilateral (MN)</td><td colspan="2">Bilateral com 802.1X (MN, AP e rede)</td></tr>
<tr><th colspan="2">Distribuição de Chaves</th><td>-</td><td colspan="2">EAP ou PSK, 4-way handshake</td></tr>
<tr><th colspan="2">Política de gestão de VIs</th><td>-</td><td>TKIP</td><td>AES-CCMP</td></tr>
<tr><th colspan="2">Cifra</th><td colspan="2">RC4</td><td>AES-CTR</td></tr>
<tr><th rowspan="2">Controlo de integridade</th><th>Cabeçalho</th><td>-</td><td>Michael</td><td rowspan="2">AES
CBC-MAC</td></tr>
<tr><th>Dados</th><td>CRC-32</td><td>CRC-32, Michael</td></tr>
</table>

Tabela 2.2: Visão global das funcionalidades de segurança com WEP, WPA e 802.11i.

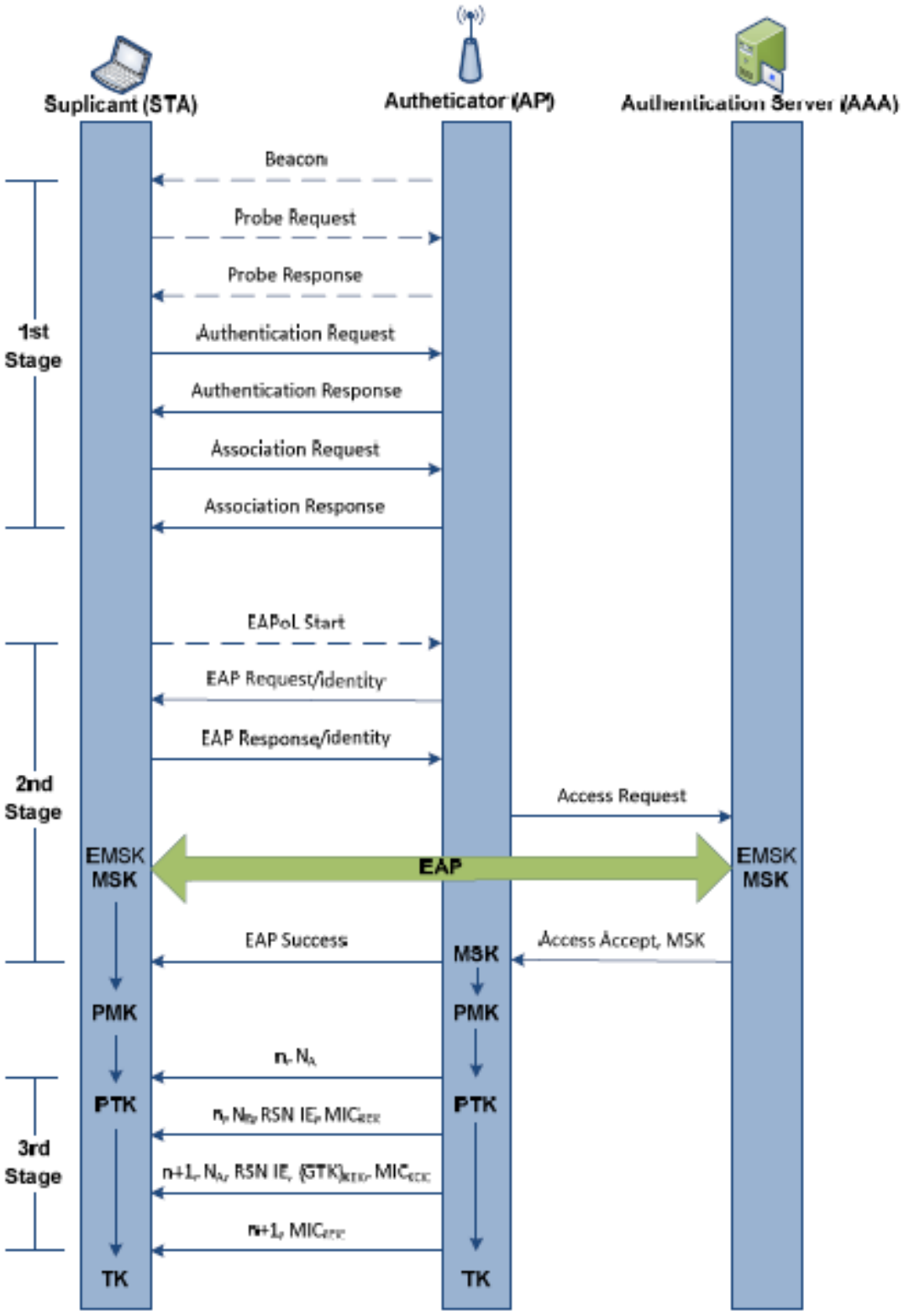

Figura 2.14: Autenticação 802.1X completa.

2.8 - Pré-autenticação em 802.11i

Uma das funcionalidades principais adicionadas ao 802.11 foi a pré-autenticação. Esta funcionalidade é extremamente relevante na abordagem do *handoff*. A complexidade adicional conferida pelo 802.11i manifesta-se em termos de latência entre a associação 802.11 e o instante em que o MN pode começar a usufruir dos recursos da rede 802.11. Se este atraso acontecer apenas uma vez no início da sessão, o atraso adicional na ordem das centenas de milissegundos pode ser aceitável. Contudo, numa situação de mobilidade, este atraso torna-se extremamente significativo e incompatível com o processo de *handoff*.

A pré-autenticação reduz a latência através do armazenamento (*caching*) de algumas chaves derivadas durante a primeira autenticação em APs vizinhos com probabilidade de serem alvo do *handoff* de um MN. Isto é conseguido realizando a autenticação nos APs candidatos a destino de *handoff* através do AP actual e do DS que geralmente interliga os APs. Esta técnica possibilita o diálogo de controlo entre o MN e APs candidatos sem interrupção do fluxo de dados a decorrer. Apesar de a latência ser diminuída, a pré-autenticação ainda introduz demasiada latência no processo de *handoff*.

Quando um MN não se encontra associado com um AP apenas pode enviar tramas de gestão. É por isso que alguns algoritmos pré-autenticam o MN via AP actual. Neste caso o MN envia tramas 802.1X encapsuladas em tramas de dados para que o AP actual possa encaminhá-las para o novo AP. Isto cria complexidade na rede porque, por questões de segurança, deverão existir associações de segurança entre APs vizinhos. Assim é necessária a criação de *Neighbour Graphs* e uma nova entidade deverá ser criada para gerir esses NGs e gerir um novo conjunto de chaves. Para tal também seria necessário algum tipo de arquitectura centralizada, como *CAPWAP* [13], para IAPP que irá criar carga adicional de tramas de gestão na rede.

2.8.1 - Passos envolvidos na pré-autenticação 802.11i

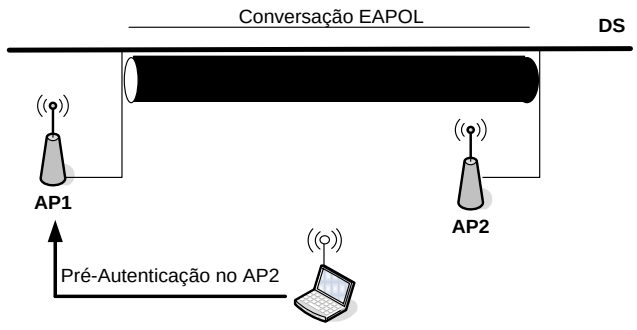

Figura 2.15: Exemplo de um MN actualmente associado ao AP1 a requerer pré-autenticação com o AP2.

O MN associa-se em primeiro lugar com o AP1 e realiza os procedimentos 802.11i normais. Após se encontrar associado, o MN toma conhecimento dos APs na vizinhança através de processos de aquisição já antes descritos (interacção via rádio ou NGs). Estes APs são colocados na lista de APs candidatos do driver da NIC sem fios. Estes candidatos são geralmente APs na mesma BSS.

O MN inicia então uma conversação EAPOL normal com cada um dos candidatos através do AP actual. Este é um passo fundamental na pré-autenticação, pois resolve o problema de não ser possível comunicar com os APs candidatos sem quebrar a associação com o AP actual. Note-se que, para um MN poder enviar tramas de dados a um AP deve estar associado a ele e o MN apenas pode associar-se a um AP de cada vez. Ao interagir com outros APs através do AP actual e do DS, o MN pode levar a cabo a conversação EAPOL normal necessária para pré-autenticar-se com outros APs de acordo com as especificações 802.11i. A conversação EAP levada a cabo em tramas EAPOL devem ter tratadas com um servidor AAA, mas esta porção pode ser feita tal como é feita normalmente, usando pacotes RADIUS como meio de transporte na rede cablada.

Cada uma destas pré-autenticações deriva uma PMK, uma para cada par MN/AP. O MN deve manter o registo sobre a relação de cada PMK com cada AP a que pertence. Tanto o MN como o AP mantêm um identificador chamado *PMK Identity* (PMKID) que lhes permite confirmar que estão a usar a PMK correcta quando iniciam a troca 4WHP.

Contudo um cliente apenas se pode pré-autenticar com APs da mesma sub-rede e envolve todas as entidades 802.11i (Suplicante, Autenticador e Servidor de Autenticação). Relativamente ao *handoff* rápido, este método ainda comporta demasiado atraso uma vez que o 4WHP tem de ser efectuado a cada transição.

Outro problema que se verifica nesta abordagem reside no facto de o MN necessitar estar associado a um AP de forma a poder efectuar pré-autenticação, desta forma os *handoff*s que requeiram pré-autenticação não podem ser executados quando o MN perde a conectividade com o seu AP de serviço actual.

2.9 - 802.11r, Transição rápida de BSS

O padrão 802.11 suporta *roaming* numa forma simples. Contudo, a velocidade de execução do *handoff* e a segurança que o padrão confere a este processo torna-o inútil quando a finalidade é conseguir uma transição de AP mantendo a ligação activa sem interrupção. O 802.11r representa uma abordagem com vista a reduzir o tempo de *handoff*. Este é ainda um padrão longe de ser ratificado quando comparado com outros padrões de complemento ao 802.11.

O 802.11r [14] define o termo *Domínio de Mobilidade* (*Mobility Domain*, MD) como sendo o conjunto de APs com capacidade de transições rápidas para os quais um MN pode transitar no momento

actual. Todos os APs neste domínio são interligados por um único DS. Para descrever MNs e APs que integrem 802.11r, o padrão define os termos *Fast Transition Enabled STA* (TSTA) e *Fast Transition Enabled AP* (TAP).

2.9.1 - BSS Pre-802.11r

Baseado nas técnicas pré-802.11r, uma transição de BSS para uma aplicação segura com suporte de QoS requer os seguintes passos:

1. Procurar APs alvo para a transição.
2. Autenticação OSA 802.11.
3. Reassociação.
4. Derivação da PTK. A complexidade deste passo varia consoante é usado *caching* de chaves, pré-autenticação ou uma nova autenticação 802.1X completa na disponibilização da PMK no novo AP. Em qualquer dos casos é necessário executar o 4WHP para derivar a PTK.
5. Controlo de admissão QoS com o novo AP.

Mesmo no melhor caso de execução destes cinco passos, o tempo total da transição de BSS irá tender para as dezenas de milissegundos ou até mais, o que irá interromper uma conversação de voz. Adicionalmente, o facto da admissão QoS ocorre apenas no final deste processo lento, o que significa que, no caso de uma recusa na admissão QoS, o processo deve ser completamente realizado deste o início para outro AP candidato.

2.9.2 - Abordagem do padrão 802.11r

A expressão "transição rápida de BSS" implica que o padrão 802.11r tenta definir procedimentos e protocolos que resultam num *handoff* 802.11. Na realidade, o 802.11r será capaz de produzir uma transição segura e rápida com suporte de QoS entre BSS que não seria possível sem 802.11r. O padrão atinge estes objectivos considerando os atrasos no processo de *handoff* que têm lugar na segurança 802.11i e no QoS 802.11e.

Os atrasos associados com a detecção da necessidade de *handoff* e a selecção do AP de destino (atraso de perscrutação da rede) bem como restabelecer os fluxos de dados das aplicações (atraso de continuação de aplicações e atraso de encaminhamento da infra-estrutura) estão além da área de trabalho do 802.11r, contudo contribuem significativamente para o atraso total do *handoff* experimentados pelo cliente. Quanto mais tempo pode ser gasto em actividades como estabelecimento do contexto de QoS e segurança durante o *handoff*, mesmo com o cumprimento completo do 802.11r, pode não ser suficiente para satisfazer as necessidades do utilizador.

O padrão 802.11r define vários termos na tentativa de impor ordem no caos da discussão dos primórdios do 802.11 centrado no *handoff* rápido e seguro. Dois dos termos mais importantes são o *domínio de mobilidade* e *primeiro contacto*. Primeiro contacto para um determinado MN é definido como sendo a sua associação inicial com o AP num MD. Os procedimentos nesse primeiro contacto são diferentes daqueles levados a cabo nas subsequentes associações desse MN no MD. Estas associações posteriores são transições rápidas de BSS, ponto principal do 802.11r.

O objectivo principal do 802.11r relaciona-se com a redução do overhead de segurança durante as transições de BSS. Os dois benefícios relativos a segurança obviamente verificados são a clarificação de como o *caching* oportunista de chaves é conseguido e a eliminação do 4WHP que tradicionalmente se seguia à reassociação. Para conseguir os efeitos desejados com o 4WHP sem o realizar é adicionada carga às quatro tramas de autenticação e associação contendo elementos de informação que transportam a informação relevante do 4WHP.

No caso do OKC (*Opportunistic Key Caching*), não são estipulados métodos pelos quais o AP obteve a informação de chaves que o MN e o AP desenvolveram durante o primeiro contacto. Em vez disso, o termo *oportunista* foi usado porque as chaves do primeiro contacto de alguma forma chegam ao AP. O *caching* oportunista de chaves assume que a informação de chaves seria disponibilizada ao AP destino tanto por algum protocolo IAPP específico do fabricante ou, no caso de arquitecturas *switch* sem fios, devido ao facto de a troca a 4 passos ter sido controlada centralmente para todos os APs no ESS e que o controlador central possuía a PMK desde o início. O facto de o processo não ter sido ditado pelo padrão resultou numa aplicação inconsciente do OKC. O 802.11r aborda este problema definindo uma nova hierarquia de chaves e o conceito de MDs tal como apresentado previamente.

Um conjunto de APs forma o MD ganhando assim acesso a uma hierarquia de chaves comum. Isto pode acontecer facilmente se o MD for desenhado para ser um grupo de APs ligados a um controlador central ou conjunto de controladores do mesmo fabricante, mas a especificação destes mecanismos está além do campo de acção do 802.11r. Este controlador é geralmente denominado *Mobility Domain Controller* (MDC) no contexto 802.11r. Um *switch* sem fios centralizado serve perfeitamente para desempenhar este papel.

Durante o *handoff*, se o AP destino anuncia a sua filiação no MD, o cliente pode associar-se com esperança que a hierarquia de chaves necessária se encontra disponível. Deve notar-se que o 802.11r não garante que a hierarquia de chaves esteja presente no momento de *handoff*, situação na qual pode ocorrer latência inesperada onde o AP terá de obter informação de quem a detém.

Quando o MN efectua a sua primeira associação e autenticação completa no MD, ganha acesso à hierarquia de chaves 802.11r. Esta hierarquia contém material PMK-R0 relativo à PMK 802.11i. Desta informação é derivada a PMK-R1 igual a uma PMK-R1 derivada pelo cliente que tenta associar-se com o AP baseado no BSSID. Este processo pode ocorrer sem recorrer ao servidor RADIUS ou qualquer outro dispositivo externo, assumindo que o AP possui a hierarquia de chaves.

A forma exacta como o AP destino obtém a informação de chaves é independente da implementação. Até ao momento não se espera que o IAPP seja ratificado. Este poderia ser a solução para a migração de chaves entre os APs. De facto, o trabalho formalizado do 802.11F, visando a ratificação do IAPP, foi rescindido. É mais provável, a esta altura, que o trabalho do grupo IETF CAPWAP seja terminado, onde é desenvolvida a comunicação entre os APs e os *switches* da infra-estrutura que os liga.

2.10 - *Caching* Oportunista de Chaves (*Opportunistic Key Caching*, OKC)

Foi apresentada previamente a pré-autenticação IEEE 802.11i. A finalidade deste mecanismo complexo é distribuir material criptográfico para APs candidatos a destino de *handoff* antes que o *handoff* ocorra. Se nos libertarmos das imposições do 802.11i e o seu objectivo ganancioso de interoperabilidade global, podem implementar-se mecanismos que vão além do padrão para concretizar esta distribuição de chaves, mecanismos esses mais simples e por vezes mais poderosos que a pré-autenticação, como é exemplo do OKC. Em particular, se uma rede implementa um dispositivo de gestão centralizada a controlar o MD de um conjunto de APs, é possível ao gestor disponibilizar a PMK 802.11i a cada AP nesse domínio de mobilidade. Tal dispositivo de gestão centralizada pode ser implementado via APs autónomos a usar IAPP específico do fabricante, mas é o *switch wireless centralizado* que apresenta a melhor oferta de soluções de *caching*. Com estes dispositivos, a PMK é transferível aos APs de baixa carga sobre o controlo do *switch* sem fios centralizado. De facto, em algumas arquitecturas de *switches* centralizados, a PMK nunca precisa ser enviada ao AP já que a cifra é feita inteiramente dentro do *switch* em vez do AP. Note-se que por vezes se usa o termo “fat” AP para enunciar o AP autónomo que efectua processamentos inteligentes e “thin” AP para enunciar o AP de baixa carga, ou seja, apenas processamento básico de serviço de rede.

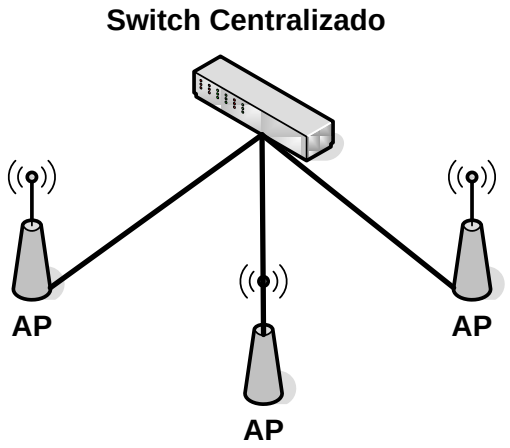

Figura 2.16: Switch Centralizado que coordena o funcionamento dos APs.

Seja qual for a forma como este *caching* das chaves seja feito, para beneficiar dele a associação com o AP deve verificar que o AP possui a PMK, que foi obtida por uma solução específica do fabricante, e portanto o cliente em associação apercebe-se que pode efectuar o 4WHP de imediato para derivação da PTK. Este processo é denominado *caching oportunista de chaves* (OKC), já que o processo não descreve como a PMK atinge o AP destino, mas apenas descreve como o MN irá oportunistamente tirar partido disso.

O OKC conta com um novo elemento de informação 802.11i, o PMKID, que foi inicialmente definido para suportar a pré-autenticação. O padrão 802.11i obriga que o IE PMKID deve ser percebido e correctamente processado por APs e MNs envolvidos. No OKC, o PMKID revela ao MN que este está associado com um *switch* sem fios, e assim, em vez de realizar pré-autenticação com APs candidatos, assume que o AP de destino conhece a PMK actual. Esta informação é codificada no PMKID, que é enviado no *Reassociation Request* na forma de IE. Se a PMK é conhecida no novo AP a autenticação fica reduzida ao 4WHP e o *handoff* é rápido. Se a PMK não é conhecida então deverá ser efectuada uma autenticação EAP completa.

Um *switch* sem fios configurado para realizar OKC não irá anunciar a capacidade de efectuar pré-autenticação. A ausência da pré-autenticação irá induzir o cliente a tentar o OKC. O pior que pode acontecer é a falha do OKC o que dará lugar a uma autenticação completa. Na prática isto não decorre desta forma num MD com capacidade de caching oportunista de chaves.

Enquanto a performance do OKC iguala a da pré-autenticação em termos de redução da troca de tramas no momento de transição, é muito superior em termos de sobrecarga de autenticação no Suplicante e no AS. O cliente de pré-autenticação antecipa o *handoff* a um determinado número de APs e completa autenticações ao estilo do 802.11X com cada um deles, inclusivamente com APs para onde o MN nunca poderia transitar.

Uma crítica comum à pré-autenticação é que esta pode mesmo aumentar a carga total de autenticação no AS comparando com a realização de uma única autenticação 802.11i completa em cada transição efectiva. O OKC é superior a ambas estas abordagens; há apenas uma única autenticação 802.11i quando o cliente entra no MD, e não haverá mais nenhuma autenticação neste estilo desde que o MN não transite para fora do MD. A única coisa que iria implicar uma nova autenticação completa dentro do MD seria a expiração da sessão AAA, cujo tempo de vida é configurável pelo gestor de rede.

2.11 - Arquitectura de Switch Centralizado Sem Fios

A implementação 802.11 clássica consiste num switch de rede cablada 802.3 com APs inteligentes a ele ligados conferindo acesso a uma rede 802.11. Estes APs são responsáveis por gerar e processar tramas de gestão 802.11 de forma a manter e reportar estatísticas de tráfego e gerir atributos de

segurança (autenticação e cifra). Na arquitectura centralizada, algumas ou todas estas funções são tratadas por um switch centralizado. A concentração da gestão da rede num só dispositivo apresenta-se uma mais-valia na resolução de problemas de coordenação entre APs. Na verdade também traz desvantagens no âmbito da escalabilidade. Um bom exemplo de problemas de coordenação de APs apresenta-se no instante do handoff de um MN entre APs.

2.11.1 – Processamento MAC

Quando nenhum processamento da camada MAC se processa ao nível do AP mas sim ao nível de uma estrutura centralizada denomina-se implementação MAC local. Todo o processamento dos pacotes, incluindo gestão de QoS e funções relativas a cifra, ocorre apenas no controlador. Como o AP não está envolvido na cifra deixa de ser necessário que o AP possua a PMK, tornando o OKC uma solução apropriada para a arquitectura MAC local.

2.11.2 - LWAPP, CAPWAP e SLAPP

Um esforço recente do IETF para generalizar o protocol de comunicação entre switches sem fios e APs foi levado pelo grupo de trabalho LWAPP (*Light Weight Access Point Protocol*). O LWAPP denomina o switch sem fios como router de acesso (*Acess Router,* AR). Apesar de vários fabricantes construírem implementações com base em drafts do LWAPP, estes nunca foram ratificados e nunca daí resultou interoperabilidade entre soluções proprietárias. Uma actividade mais recente do IETF com a mesma visão é levada a cabo actualmente pelo grupo de trabalho CAPWAP (*Control And Provisioning of Wireless Access Points*). Neste grupo o switch wireless é chamado Controlador de Acesso (*Access Controller, AC*). O CAPWAP seleccionou o LWAPP como base na generalização do protocolo de comunicação entre switches sem fios e APs. Outro empenho, o SLAPP (*Simple Access Point Protocol*) tencionava ser menos ambicioso que os antecessores, contudo parece não ter continuado a sua actividade até agora.

O CAPWAP será apresentado explicado e no capítulo 3.

2.12 - 802.11k

O padrão 802.11k [15] encontra-se em desenvolvimento e irá permitir transições transparentes entre BSS num ambiente sem fios. O padrão fornece informação útil para a descoberta do melhor AP disponível na vizinhança do MN.

2.12.1 - O Padrão 802.11k

O padrão define uma série de pedidos e relatórios de determinadas medidas que detalham estatísticas de cliente sobre as camadas 1 e 2. Geralmente os APs pedem aos clientes relatórios sobre a sua ligação, mas em alguns casos os clientes poderão pedir informação aos APs.

O objectivo do 802.11k consiste no melhoramento da distribuição de tráfego na rede sem fios. Nesta rede o MN simplesmente se associa ao AP que oferece a potência de sinal superior. Dependendo do número de clientes e da sua localização geográfica, este método poderá conduzir à exigência excessiva de recursos de um único AP e utilização escassa de outros resultando na degradação da performance geral da rede. Numa rede onde seja aplicado o 802.11k, se o AP com melhor sinal se encontra carregado à sua capacidade máxima, um MN transfere a sua associação para outro AP subutilizado que apresenta naquele ponto geográfico menor potência de sinal. Apesar da potência de sinal ser inferior, o desempenho global é superior graças à utilização mais eficiente dos recursos da rede.

Antes da transição de AP, a seguinte linha de processos tem lugar: o AP determina que o MN se afasta geograficamente de si. Nesta sequência o MN é informado pelo seu AP actual que deve preparar-se para sofrer uma transição para outro AP. Com esta notificação o MN efectua o pedido de uma lista de APs na sua vizinhança para os quais a probabilidade de *handoff* é superior. Por fim o MN efectiva o *handoff* levando em conta a informação recebida do AP.

2.13 - Servidor HOKEY (*Handover Keying*)

Arquitectura proposta pelo grupo de trabalho HOKEYWG (HOKEY Working Group). Este grupo empenha os esforços em conseguir um *handoff* rápido eliminando execuções repetidas de autenticações EAP a cada vez que o *handoff* ocorre uma vez que é usado o mesmo servidor de autenticação. A ideia base do protocolo consiste na utilização de chaves não expiradas resultantes da comunicação EAP evitando uma nova e extensa sessão EAP com o servidor de autenticação.

O HOKEYWG usa a EMSK (Extended Master Session Key) como a base da hierarquia de chaves usada para implementar os serviços de segurança requisitados. Também fornece raciocínios e recomendações para a derivação das chaves consequentes. As soluções especificadas pelo HOKEYWG enquadram-se em diversas categorias, baseadas em timing e mecanismo. A autenticação e gestão de chaves podem ter lugar antes do handoff, quando a latência é menos crítica. Soluções devem reduzir ou eliminar o número de intercomunicações com servidores de autenticação e estas soluções devem evitar executar comunicações EAP repetidas. Isto pode ser conseguido conferindo novos mecanismos de criação de chaves combinados com mecanismos de entrega deste material criptográfico às entidades apropriadas.

Capítulo 3

Trabalhos Relacionados

O nosso trabalho visa optimizar a mobilidade de um cliente no âmbito de redes sem fios IEEE 802.11. Assim, nesta secção, são explorados os parâmetros que afectam o MN enquanto este se movimenta geograficamente através da área de abrangência de uma rede servida por APs pertencentes a um domínio. Apresentam-se então as métricas que podem ser exploradas para efectuar decisões que afectam a mobilidade, isto é, decisões de *handoff*, quer relativamente à escolha do AP mais apropriado a ser o alvo da transição, quer o instante mais vantajoso para o fazer.

Os parâmetros e métricas são as características que cada dispositivo apresenta influenciadas pelas limitações da sua própria implementação (características de hardware), pelo meio em que se encontram incluídos (densidade de dispositivos a partilhar o meio, atenuações e interferências proporcionados por obstáculos físicos e electromagnéticos) e pelos serviços que cada um tem a capacidade de fornecer/usufruir (protocolos configurados, capacidades de comunicação e compatibilidade entre pares de dispositivos).

Apresentam-se também neste capítulo trabalhos de diversos autores que abordam as mesmas questões que são tratadas neste trabalho e tentam, de diversas formas, melhorar aspectos de mobilidade implementando diferentes mecanismos de *handoff* e reforçando diversas políticas de gestão do processo. Esta segunda secção deste capítulo encontra-se organizada em duas partes: na primeira secção são expostas soluções propostas para redução da latência de *handoff* mediante a redução da latência resultante do processo de perscrutação de APs próximos, bem como outras técnicas que não incorporam melhorias em protocolos de segurança. Na segunda secção são expostas técnicas, políticas e propostas de *handoff* que visam reduzir o impacto da latência proporcionada pelo estabelecimento dos mecanismos de segurança durante a transição do MN, nomeadamente a redução da latência conferida pelo processo de autenticação mútua 802.1X, que se relaciona directamente com a solução que nós desenvolvemos neste trabalho.

3.1 - Métricas e Atributos de Rede

O *handoff* ocorre sempre que um MN precisa de mudar a sua associação de uma BSS para outra. Não está mencionado nas especificações do IEEE 802.11 qual a técnica usada para decidir quando efectuar o *handoff*. Contudo, a forma mais comum, seguida pela maior parte dos fabricantes, consiste em iniciar a fase de descoberta e fases consequentes sempre que a força do sinal recebida desce abaixo de um limiar predefinido. Porém, muitas abordagens exploram outras propriedades da comunicação sem fios. Essas propriedades mensuráveis, isto é, que podem ser medidas ou contabilizadas, são apresentadas de seguida. Também se apresentam os parâmetros relevantes que influenciam ou podem ser influenciados pela abordagem que for tomada e devem ser analisados e levados em conta.

3.1.1 - RSSI (métrica)

O RSSI recebido no MN indica a potência do sinal recebido ao nível do cliente no ponto geográfico em que este se encontra quando recebe os pacotes provenientes do AP. Quanto mais perto o MN se encontra do AP, melhor o RSSI medido nesse ponto. Adicionalmente pode ser considerada também a potência de sinal recebido ao nível do AP medido na recepção dos pacotes enviados pelo MN.

Factores externos interferem na potência de sinal recebida, nomeadamente atenuação no ar, obstáculos tais como portas, janelas ou até pessoas onde as ondas electromagnéticas são reflectidas, refractadas e atenuadas. Deter informação sobre o RSSI em ambos os sentidos do canal de comunicação demonstra-se uma mais-valia quando o objectivo é melhorar a performance do sistema e criar políticas de decisão de *handoff*.

Um canal de comunicação sem fios pode apresentar diferentes RSSIs nos diferentes sentidos de comunicação, o que afecta directamente a performance da ligação em ambos os sentidos da comunicação entre o AP e o MN. Para soluções que atribuam ao MN a responsabilidade de gerir o *handoff* e que manifestam interesse na medição do RSSI em ambos os sentidos, propõe-se que essa informação medida a nível do AP seja enviada na forma de IE em tramas enviadas ao MN (já que o MN já detém a informação sobre o RSSI no sentido inverso). Aconselha-se adicionalmente que a trama escolhida seja especificamente um Probe Response, pois esta trama pode ser requisitada pelo MN durante a fase de descoberta sem necessidade de existir nenhuma associação com o AP a ser indagado.

Estudos recentes demonstram que uma abordagem simples de associação baseado apenas no RSSI conduz a uma utilização ineficiente da rede em termos de distribuição de utilizadores pelos diferentes APs e por desprezar outros parâmetros que não são levados em consideração [2], [16], [17].

3.1.2 - SNR (Signal-to-Noise Ratio) (métrica)

Este parâmetro consiste na divisão da potência de sinal pela potência do ruído observado num determinado ponto da comunicação. Um valor elevado de SNR traduz-se em boa qualidade de sinal. O ruído é produzido por diversas fontes, ruído electrónico produzido por dispositivos eléctricos, ruído causado por vibrações, humidade, vento, e outros factores. Abordagens ao *handoff* podem usar esta métrica no estabelecimento de políticas de mobilidade. Mais uma vez, se o MN for a entidade responsável pela gestão do processo, poderá receber a medição do SNR ao nível do AP num IE adicionado a uma trama tal como o método descrito na alínea anterior. Como este factor varia com a distância, interferência de outros dispositivos, obstáculos e outros factores, o SNR do canal deve ser medido durante um período de teste e não pontualmente (valor instantâneo) para que se possa avaliar este factor.

3.1.3 - LB (Largura de Banda) (métrica)

A largura de banda disponível para uso de um dispositivo móvel é um factor importante que deve ser levado em conta quando é feita a avaliação do AP para o qual o MN vai transferir o estado de associação. Este parâmetro é influenciado pelo número de dispositivos apensos ao AP e pelo tráfego que estes injectam no meio. Se um AP suporta o serviço de um número elevado de utilizadores, a probabilidade de colisão aumenta. A largura de banda disponível é limitada pelo grau de saturação imposto pelos dispositivos móveis quando a sua demanda e injecção de dados na rede aumenta. Desta forma estes valores devem ser comunicados pelo AP ao MN, uma vez que o AP pode ter inúmeros MNs a si associados mas com baixo débito dos recursos da rede ou poucos MNs associados com elevado débito.

Torna-se relevante a avaliação destes valores para determinar o alvo do *handoff* de um MN, como também é relevante que a rede se organize de forma a não sobrecarregar um AP e subutilizar outros. Este é um problema abordado pelas recomendações do grupo de trabalho 802.11T. Convém também referir que este parâmetro é ainda afectado pelos diferentes protocolos que conferem overhead adicional que debita LB para o seu funcionamento. Na avaliação das características da LB podem diferenciar-se capacidades de Uplink e Downlink. As taxas de transmissão podem variar dependendo do sentido da comunicação. Isto depende da taxa de download ou upload verificada a nível do AP. Ao nível do MN, um AP pode parecer apropriado quando considerada a oferta de largura de banda para download, mas após a associação o MN apercebe-se que a taxa de upload é bastante limitada devido ao uso excessivo de outros MNs. Considerar a oferta de LB em ambos os sentidos da comunicação demonstra-se importante através do trabalho realizado em [18].

3.1.4 - Perda de Pacotes (métrica)

Perda de pacotes representa o número de pacotes que não podem ser recebidos no seu destino. Este problema pode ter diversas fontes tais como falha da ligação, congestionamento elevado da rede que provoca overload dos buffers, degradação e atenuação do sinal, pacotes corrompidos ou falhas de hardware.

A perda de pacotes afecta a performance geral da rede. A avaliação da qualidade da ligação pode basear-se na taxa de perda de *Beacons*. Quando um MN se encontra associado a um AP, mesmo em *Power Save Mode*, mantém-se atento à difusão de Beacons desse AP. A perda de Beacons é um indicador da qualidade da ligação, quando esta perda acontece o MN não consegue determinar a razão para tal pois diversas justificações são possíveis (interferência ou atenuação do sinal, demasiada carga no AP que adia o envio do Beacon, etc.). A perda de um determinado número de *Beacons* consecutivos revela um problema de comunicação com o AP actualmente a servir o MN, factor este a levar em consideração para a implementação de políticas de mobilidade.

3.1.5 - Atraso, Latência ou Lag (métrica)

Várias componentes de atraso podem ser contabilizadas quando se aborda esta métrica. O atraso quantifica o tempo dispendido na transmissão de um pacote entre um par de nós da rede em comunicação. Este atraso pode ser significativo ao ponto de impedir o funcionamento de determinadas aplicações sensíveis, como por exemplo o streaming de dados ou aplicações de tempo real como vídeo-conferência. Atrasos elevados traduzem-se na degradação da qualidade de imagem ou pausa entre frames consecutivos. Uma rede 802.11 preparada para suportar estas aplicações deve garantir limites de atraso de transmissão, bem como limites de atraso entre pacotes consecutivos.

No caso de transmissão de vídeo o limite máximo de atraso permitido é de 200ms enquanto, no caso de vídeo-conferência o limite máximo é de 100ms e uma chamada VoIP permite atrasos máximos na ordem dos 50ms, segundo Ganesh Venkatesan em [19]. Por outro lado, quando se carregam/descarregam dados, por exemplo, durante um download de um ficheiro, não se verifica a relevância do atraso dos pacotes. O atraso é influenciado por diferentes factores como proximidade física do MN ao AP, protocolos a correr na comunicação, existência ou não de políticas de QoS, LB disponível, etc.

Este aspecto deve ser levado em consideração quando se implementam políticas de mobilidade, quer seja diferenciado o serviço que o MN prevê que irá necessitar aquando a selecção de um AP apropriado, quer sejam implementadas políticas de forma geral para todos os MNs de forma a proporcionar o melhor serviço possível para todos deles, independentemente do tipo de tráfego que vá ser transaccionado na rede.

3.1.6 - Jitter (métrica)

Jitter pode ser simplesmente explicado como sendo a variação do atraso, isto é, diferentes pacotes demoram tempos diferentes a serem transmitidos, o que resulta na alteração da ordem de recepção de pacotes durante a comunicação e impossibilidade de prever a recepção correcta do próximo pacote da cadeia. Aplicações VoIP e de videoconferência apresentam performance bastante degradada na presença de jitter. Este é um factor a levar em conta na elaboração de uma solução de mobilidade, pelo que o seu controlo é imprescindível, essencialmente quando se pretendem usar aplicações de VoIP nessa comunicação sem fios. Segundo [19], comunicações VoIP possuem o limite estipulado de 50ms para o jitter.

3.1.7 - Débito de Energia (métrica)

Esta é uma característica apenas referente a cada MN. Um dispositivo móvel dissipa energia armazenada numa bateria, sem necessidade de estar ligado à rede eléctrica. Esta é a propriedade que cria o conceito de dispositivo móvel e que permite a exploração da mobilidade geográfica. No que diz respeito a este trabalho, o débito de energia é um parâmetro a ser considerado quando se tenciona melhorar a experiência de mobilidade através da adição de novos dispositivos a funcionar ao nível do MN.

Por exemplo, a abordagem dos autores do trabalho referenciado em [20] usa duas NICs no MN em vez de uma com o objectivo de implementar uma solução que melhora a mobilidade mediante o melhoramento do processo de perscrutação do meio, contudo o débito energético aumenta. Outros trabalhos sugerem o uso de dispositivos de localização (ex: GPS). Estes dispositivos adicionais resultam no aumento do débito energético, indesejável em comunicações móveis.

3.1.8 - Escalabilidade (atributo)

Escalabilidade representa a capacidade de adicionar novos utilizadores a um AP. Novos utilizadores implicam concorrência acrescida ao meio, pelo que este deverá ser limitado para garantir uma qualidade de serviço aceitável para cada uma das associações. Apesar de um AP permitir a associação de novos dispositivos não significa que haverá garantias de uma determinada performance que o MN possa necessitar. Por outro lado, o AP pode já servir todos os MNs que consiga mas, considerando o débito imposto para estes MNs, poderia ainda haver espaço para novos dispositivos em termos de largura de banda disponível. Não existe nada no padrão 802.11 que resolva este problema, pelo que cada AP se limita a aceitar a associação de MNs autenticados degradando o serviço sem contenção de danos na comunicação.

3.1.9 - Atributos da Rede de Destino (atributo)

A rede servida por um AP ao alcance de um MN pode não fornecer serviços considerados importantes para a comunicação actual do MN. Por exemplo, se o MN demonstrar interesse por um ponto de acesso seguro com mecanismos de autenticação, pode ter de recusar um AP avaliado como sendo melhor que outros em termos de performance de comunicação mas invalidar a possibilidade de se associar a este pela não existência dos mecanismos de segurança.

No caso específico da segurança informática, esta tem como objectivo garantir privacidade do utilizador, integridade dos dados e controlo de acesso a uma rede. Uma rede sem segurança ou com segurança pobremente elaborada coloca em risco os utilizadores dessa mesma rede. A avaliação da capacidade da rede para proteger os MNs é uma tarefa que deve ser tomada em atenção nas políticas de mobilidade. O AP divulga essas características de segurança nos *beacons* e nos P*robe Responses*, permitindo ao MN obter essa informação e efectuar as suas decisões mediante esse conhecimento.

3.1.10 - Hardware Específico utilizado (atributo)

O hardware sem fios utilizado (cliente e AP) afecta a latência de *handoff*. A experiência levada a cabo em [2] demonstra diferentes atrasos consoante o uso de diferentes pares MN/AP. As condições desta experiência são as seguintes: diferentes dispositivos móveis (em termos de fabricante específico, Cisco, Lucent e ZoomAir) movimentam-se ao longo de um meio servido exclusivamente por APs Lucent numa primeira experiência e Cisco numa segunda fase da experiência. É medido o tempo de *handoff* da camada MAC de cada MN enquanto este transfere a associação ao longo dos APs a ser testados. A autenticação foi Open System e os valores apresentados representam uma média dos tempos de handoff medidos em cada etapa da experiência. Estes resultados apresentam-se de seguida:

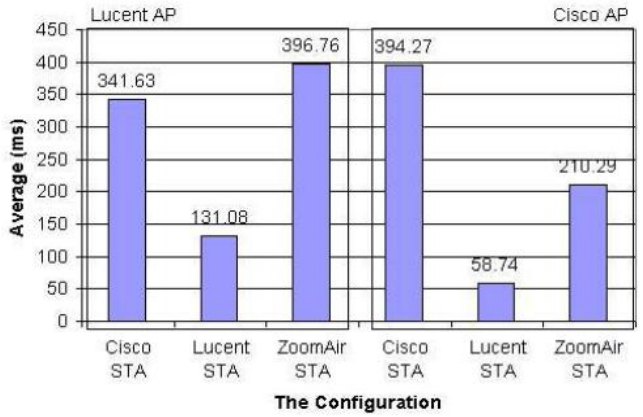

Figura 3.1: Valores médios das latências de *handoff* entre diferentes pares MN/AP.

Verificam-se grandes variações na latência de *handoff* quando se emparelham diferentes pares de MNs/APs. Além das variações na latência entre diferentes configurações (pares cliente/AP), esta experiência também demonstra variações significantes entre *handoff*s sucessivos com a mesma configuração. O trabalho feito em [2] também refere a existência de diferentes sequências de mensagens entre NICs de diferentes fabricantes. Por exemplo, placas ZoomAir Prism2 enviam a mensagem de reassociação antes da mensagem de autenticação, que é enviada como resposta a uma mensagem de desautenticação recebida do AP. Esta sequência de mensagens viola o padrão 802.11.

3.2 - O *Handoff* 802.11

O uso de um threshold para disparar a fase de descoberta é o processo mais difundido nas implementações actuais. Este método representa uma limitação na medida em que a degradação do sinal necessária para disparar esta fase é de tal ordem que não pode haver mais comunicação, ou já não havia durante algum tempo. O momento em que a fase de descoberta do processo de *handoff* deve ser efectuada é um parâmetro importante a ser analisado.

Várias opções podem ser adoptadas além desta abordagem apresentada para decidir quando efectuar a perscrutação do meio:

- Pode ser executada consecutivamente desde o início da comunicação em intervalos específicos;

- Em períodos de baixa taxa de comunicação;

- Quando o RSSI decaí abaixo de um determinado valor. Esse threshold pode ser controlado para que a fase de descoberta seja feita antes que a comunicação se degrade demasiadamente

- Utilizar um método que não necessite que esta fase tenha lugar, tal como atribuir ao AP a responsabilidade de tratar do *handoff* dos MNs a ele associados a cada momento.

- Não efectuar nenhuma procura e receber a informação relevante fornecida por uma entidade específica para o efeito;

Existem diversos factores que se traduzem em atraso de *handoff* nas implementações desenhadas actualmente. Existe uma certa dificuldade na apresentação de valores médios para os atrasos das diferentes fases de *handoff* em redes 802.11 pois esses atrasos dependem do contexto global da implementação.

3.2.1 - Atrasos característicos das fases do *handoff*

No artigo [3] o autor apresenta a latência sofrida nas diversas fases do *handoff* 802.11. Esses valores são apresentados de seguida.

Camada OSI	Item	Tempo (ms)
Camada 2	Scan 802.11 (passivo) Scan 802.11 (activo) Assoc/reassoc 802.11 (sem IAPP) Assoc/reassoc 802.11 (com IAPP) Autenticação 802.1X (completa) Autenticação 802.1X (com pré-autenticação) *Handoff* Rápido (apenas 4WHP)	0 (cache), 1000 (espera do beacon) 40 a 300 2 40 1000 250 60
Camada 3	DHCPv4 RS/RA inicial Espera de RA consequente DAD (completo) DAD Optimísta	1000 5 1500 1000 0
Camada 4	Ajuste de parâmetros TCP (status quo)	5000(802.11/CDMA) – 2000(802.11/GPRS)
Melhor Caso	Melhores casos	150
Caso Médio	6to4, RR, Scan Activo	1300
Sem TCP, auth EAP competa, IAPP, DHCPv4		2500

Tabela 3.1: Latência das fases 802.11 [3].

Apresentam-se agora os atrasos que cada uma destas fases implica e os problemas que cada fase ostenta:

• **Fase de Descoberta:** O maior problema verificado nesta fase revela-se quando se verifica que quanto mais tempo é gasto na perscrutação de outros canais, menos tempo é aplicado na troca de dados no canal actualmente utilizado pelo MN, reflectindo-se no aumento da probabilidade de perda de tramas. Assim, qualquer solução de *handoff* que concentre os esforços no aumento da quantidade de informação útil por unidade de tempo gasto fora do canal actual de comunicação irá beneficiar a performance do throughput global.

Na Aquisição passiva o atraso de scan pode ser determinado mediante a seguinte equação:

$$Atraso\ de\ scan\ passivo = Número\ de\ Canais\ .\ Intervalo\ Máximo\ entre\ tramas\ Beacon$$

O cliente permanece cerca de 100ms em cada canal para esperar a chegada dos beacons, o que se traduz, no pior caso, numa latência total de cerca de 1.1s em redes 802.11b/g, ou superior ainda se a placa sem fios suportar 802.11a. Isto representa um grande problema quando a finalidade é conseguir um *handoff* rápido. Outro problema pode ser encontrado a nível do AP, quando este não gera beacons a cada 100ms devido a uma elevada carga de rede ou atrasos de fila de espera no instante em que o MN efectua a escuta.

Quanto à aquisição activa, o valor do atraso médio sofrido pode ser calculado com a equação seguinte.

$$Atraso\ de\ scan\ activo = \sum_{c=1}^{NumCanais} \big(1 - P(c)\big).Min + P(c).Max$$

Onde **P(c)** é a probabilidade de um ou mais APs a operar no canal **c**. Para valores padrão de 1ms para o *MinChannelTime* e 11ms para o *MaxChannelTime*, a latência deveria incidir idealmente no intervalo entre os 11ms e os 110ms para o 802.11b. Contudo é necessário ter em conta o tempo de mudança de canal (*channel switch delay*) que reflecte o tempo característico de cada NIC de mudança de frequência, re-sincronização, e início da desmodulação dos pacotes recebidos. Este tempo ronda os 19ms (12ms para mudar de canal e 7ms para re-sincronização), mais uma vez, dependendo da NIC em questão. Como este atraso existe por canal, existe então uma carga considerável desta componente no atraso da perscrutação activa. Outra aproximação para o atraso sofrido por um cliente numa fase de perscrutação pode ser limitada pela seguinte expressão:

$$N \times Tb \leq t \leq N \times Tt$$

onde **N** é o número de canais, **Tb** o *MinChannelTime* e **Tt** o *MaxChannelTime*.

Este processo é mais rápido que a modalidade passiva, porém introduz mais carga de tráfego na rede. Considerando uma elevada densidade de APs na rede, o tempo de espera em cada canal será frequentemente o MaxChannelTime, o que resulta em atrasos de 50ms a 360ms para placas a operar com 802.11b/g, ou superiores se a placa sem fios suportar 802.11a.

• **Atraso de Reassociação:** Este atraso é o tempo necessário para completar a associação com o novo AP. Pode envolver a troca de tramas de autenticação e associação. O início da fase de reassociação com o novo AP marca o fim definitivo da passagem de dados de aplicação pelo AP anterior. Como apresentado na tabela 3.1, o atraso desta fase é de 2ms sem uso do IAPP e 40ms com o uso do IAPP.

• **Atraso de Autenticação:** Este atraso consiste no tempo de conversação entre o MN e o AP. Dependendo da versão de segurança a ser usada na rede, esta troca pode incluir diversas tramas. Teoricamente podem ser trocadas centenas de tramas se for necessária uma comunicação EAP com o servidor de autenticação em versões avançadas de autenticação. Na comunicação 802.1X verifica-se um valor médio de 1 segundo de atraso para concluir a autenticação mútua do MN e AP.

• **Atraso da gestão de chaves:** Tempo gasto pela troca 4WHP de tramas de gestão de chaves usadas para derivar as chaves de sessão que serão usadas para proteger a ligação sem fios.

• **Atraso de Retoma de Aplicação:** Assim que os intervenientes sejam programados com as chaves derivadas durante a fase de gestão de chaves, existe um atraso interno adicional quando os drivers enviam um evento *LINK UP* aos protocolos de níveis superiores e antes de estes reagirem e retomarem o tráfego de aplicação. MNs actuais enviam uma trama *ARP* ao *default gateway* para determinar se a sub-rede IP mudou. Na eventualidade dessa mudança de sub-rede, o atraso envolvido pelo uso do DHCP para obtenção de um novo endereço IP faz parte do atraso de retoma de aplicação.

• **Atraso de Routing da Infra-Estrutura:** Esta última fase reúne qualquer período que pode ter lugar na infra-estrutura depois do AP e do MN estarem prontos para retomar o fluxo de dados. No caso de um *handoff* local, apesar de o AP e MN estarem completamente preparados para retomar a comunicação, para que o tráfego destinado ao MN lhe seja entregue, terá de ser encaminhado pelo novo AP. Até que a infra-estrutura reaja a esse movimento, existe ainda uma perda de conectividade do ponto de vista das aplicações que esperam que os protocolos de switching estabeleçam a nova spanning-tree.

3.2.2 - A Componente Mais Significativa

A pesquisa feita em [2] demonstra que o tempo da fase de descoberta é a componente dominante no processo de *handoff*. Este atraso representa 90% de todo o processo, independentemente da combinação de clientes e APs de diferentes fabricantes. Mesmo quanto ao número de mensagens trocadas entre os intervenientes do *handoff*, as mensagens relativas à fase de descoberta representam 80% da totalidade. Leva assim a concluir que qualquer esquema de *handoff* que consiga armazenar ou deduzir informação relativa aos APs da rede sem ter de efectuar uma perscrutação activa completa, claramente beneficia da redução do custo que este atraso representa em todo o processo.

3.2.3 - Requisitos de Segurança no Processo de *Handoff*

São enunciados nesta secção quais os requisitos de segurança de uma rede 802.11 durante o processo de *handoff* [21]:

1. O atraso de *handoff* deve ser reduzido tanto para transições intra como para inter-domínio.
2. O nível de segurança não deve ser afectado e a privacidade do cliente deve ser protegida.
3. O procedimento deve impedir a transição para APs falsos.
4. Devem ser evitados os ataques MAN-in-the-Midlle (MitM) e roubo de sessão.
5. No *handoff* inter-domínio, a privacidade de uma rede não deve ser desvendada numa outra.

6. Os requisitos estabelecidos pelo IEEE 802.11i devem ser respeitados.
    a. Deve haver autenticação mútua e uma nova derivação de chaves em cada AP.
    b. A autenticação mútua não deve permitir o ataque MitM.

3.3 – Optimização do Tempo de Perscrutação

Na fase de descoberta é feita a perscrutação do meio circundante ao alcance do MN, nesta fase o MN encontra os dispositivos que disponibilizam o acesso à rede, ou seja, os pontos de acesso, ou APs.

É oportuno diferenciar dois tipos de perscrutação que uma abordagem pode adoptar na gestão deste processo. A *Perscrutação Forçada* (seja feita de forma activa ou passiva) tem lugar quando a qualidade da ligação decai até níveis inconsistentes com uma boa comunicação entre o MN e o AP. Nesta versão o MN detecta um mau serviço devido à degradação de diversos factores ou mesmo resultando do retiro do AP da rede, o tempo de perscrutação aparece então adicionado ao tempo de *handoff*. Outro tipo chama-se *Perscrutação Proactiva*, onde o MN decide proactivamente realizar a pesquisa do meio antes que o processo de *handoff* propriamente dito tenha lugar. A obtenção da informação tem lugar durante a comunicação normal enquanto a performance da ligação com o AP actual se apresenta aceitável.

A pesquisa da vizinhança pode ser feita de forma passiva, onde o MN se limita a receber e processar os beacons que os APs emitem periodicamente, ou activa, na qual o MN envia a pergunta, o *Probe Request*, esperando receber uma resposta dos APs no seu alcance, o *Probe Response*. Este processo é indispensável para recolha de informação por parte do MN quando não existe uma base de dados que disponibiliza esta informação, como por exemplo Neighbour Graphs.

Segundo o trabalho desenvolvido por Mishra et. al. em [2], o tempo de perscrutação forçada efectuado de forma activa feito durante a fase de *handoff* é a componente de atraso dominante do processo representando 90% desse tempo.

Os trabalhos apresentados de seguida visam reduzir a latência da fase de perscrutação.

Apresentam-se de seguida trabalhos que partilham o objectivo de reduzir o tempo de *handoff* através do manuseamento da fase de perscrutação. O título de cada abordagem introduz a explicação de cada uma delas.

3.3.1 - Using Smart Triggers for Improved User Performance in 802.11 Wireless Networks [22]

Actualmente a maior parte dos algoritmos de *handoff* em redes 802.11 são reactivos, ou seja, somente quando a qualidade da ligação se degrada substancialmente é disparado o processo de descoberta do meio e fases consequentes do *handoff*, e apenas se baseiam na qualidade instantânea do sinal do AP para decidir para qual vão mudar. Este processo demora certa de 2 segundos, o que não é sustentável para aplicações tais como VoIP.

Esta abordagem apoia que um cliente não deve esperar pela perda ou degradação excessiva de performance da comunicação, o processo deve ser proactivo na procura de um AP alternativo. Para tal é sugerida a monitorização contínua das ligações sem fios. Um MN mede a intensidade do beacon de todos os APs que operam no canal actual e os respectivos canais sobrepostos no espectro electromagnético tal como descrito no IEEE 802.11b/g, baseando a decisão de *handoff* na tendência da variação da qualidade de sinal a curto e longo termo.

O driver do cliente monitoriza continuamente a performance dos APs a operarem no canal actual e todos os APs a operar nos canais sobrepostos a este medindo a força de sinal dos beacons. Tais implementações implicam alterações computacionais triviais. Por conseguinte, como o firmware tenta descodificar todas as tramas recebidas por omissão, expor todos os beacons ao driver não incorre numa sobrecarga adicional. Assim, a descoberta de APs tem atrasos nulos proporcionando ao cliente medidas para selecção do próximo AP a ser utilizado. Quando não é encontrado um AP mais adequado no canal actual ou canais adjacentes e sobrepostos, o cliente retoma a fase de descoberta em todos os outros canais. Este scan geral é usado somente como último recurso, quando os resultados do scan proposto não obtêm soluções propícias ao *handoff*.

Se o Power Save Mode (PSM) estiver activo, haverá períodos em que a aquisição dos beacons nos canais não pode ser feita, o impacto do PSM não é estudado neste trabalho). Outra desvantagem é experimentada quando esta abordagem é utilizada em 802.11a pois nesta arquitectura não existem canais sobrepostos.

3.3.2 - Multimedia Ready *Handoff* Scheme for 802.11 Networks [23]

Nesta abordagem é separada a fase de perscrutação do meio da execução do *handoff*. É efectuada a perscrutação durante a troca normal de dados, quebrando o processo em fases mais curtas. O cliente usa um mecanismo de scan que corre em segundo plano. Neste scan de segundo plano o cliente envia uma trama de PSM (Power Save Mode) ao AP actual e entretanto muda para outros canais para fazer a sondagem. Quando termina volta ao canal actual. O AP irá guardar no buffer as tramas destinadas ao cliente enquanto este está em PSM.

Para impedir atrasos elevados o cliente apenas irá realizar a perscrutação em segundo plano num canal de cada vez. O scan de um canal demora no máximo 16ms (contando com a troca de canal (≈4ms), envio do P*robe Request*, espera e recepção do P*robeRresponse* (≈8ms), mudança de novo para o canal de partida). Desta forma o tempo de *handoff* reduz-se à mudança para um AP mais adequado, sendo o tempo de perscrutação eliminado por completo do processo de *handoff*. Após mudar para o novo AP, o cliente renova a lista de canais a serem sondados pelo método de scan em segundo plano e inicia o processo de perscrutação após um intervalo de tempo escolhido e iterado sobre todos os canais integrantes da lista.

Se nenhuma resposta é ouvida no canal após o tempo de espera (*MinChannelTime*), o canal é removido da lista, um novo canal é seleccionado para scan e o cliente retorna ao canal actual de funcionamento com um atraso de mudança de canal de ≈4ms. O tempo de atraso entre dois pacotes de uma aplicação VoIP (*Inter-Arrival Time*, IAT) é aproximadamente de 20ms, contando portanto que o processo de scan de um canal demora 16ms, é possível realizar o scan de um canal durante esse IAT. A lista de canais a ser sondados fica reduzida aos canais que obtiveram resposta, e futuras perscrutações serão efectuados apenas sobre estes canais. Para evitar perdas de pacotes ocasionais, o AP irá guardar em buffer os pacotes destinados ao cliente que se encontra de momento em PSM. Para evitar a perda de beacons no canal actual, não é efectuada perscrutação no instante em que esses beacons são esperados.

Para implementar esta solução apenas é necessário adicionar ao cliente uma aplicação que gere a descoberta da vizinhança sem efectuar alterações à infra-estrutura nem aos padrões existentes.

Esta abordagem requer sincronismo para que a perscrutação não seja feita no momento de recepção de um pacote e apenas durante o IAT. A informação resultante da perscrutação é guardada em cache na aplicação e consiste nos valores do RSSI medidos. Em contrapartida, apenas se baseia a decisão do AP mais apropriado para *handoff* com base em informação do RSSI e não tem em conta carga do canal ou do próprio AP.

A decisão de início de *handoff* é tomada pela aplicação criada. O *handoff* é disparado assim que o RSSI do canal actual desce abaixo de um threshold pré definido. Não é apresentada qualquer solução para tratar da transferência ou criação do material de segurança necessário em redes que implementam segurança 802.11i. O ponto forte deste trabalho, a pesquisa da vizinhança nos intervalos IAT, é abalado pelo facto de os testes deste trabalho terem sido feitos sobre uma aplicação VoIP.

Apesar da boa intenção na escolha desta aplicação, devido ao facto de ser uma aplicação extremamente sensível ao atraso de *handoff* (que deve ser inferior a 50ms), as características deste tráfego invalida os resultados para outros tipos de tráfego. O tráfego VoIP é constante e contínuo, sendo fácil ter uma janela de tempo entre pacotes para efectuar determinadas operações, porém outros tráfegos não são constantes nem contínuos, pelo que é perdida a janela oportuna para efectuar as devidas tarefas de perscrutação propostas nesta abordagem. Assim esta solução

apresenta-se limitada a tecnologias VoIP ou aplicações que apresentem tráfego com características semelhantes.

3.3.3 - SYNCSCAN: Practical Fast *Handoff* for 802.11 Infrastructure Networks [24]

Esta abordagem propõe uma técnica de baixo custo para seguir continuamente as BSSs na vizinhança sincronizando pequenos períodos de escuta de transmissões de cada BSS no cliente. Esta solução trata os problemas de *handoff* com uma simples e única preocupação: como monitorizar continuamente a proximidade de APs na sua vizinhança. Tenta substituir a elevada carga da descoberta activa de APs por um processo contínuo que faz a monitorização passiva de outros canais pela presença de APs. A potencial disrupção da mudança de canal é minimizada sincronizando pequenos períodos de escuta no cliente com transmissões periódicas dos APs.

APs emitem *beacons* periódicos para se identificarem a potenciais clientes e sincronizar informação de estado com clientes actualmente associados. Como o intervalo de envio é configurável este pode ser organizado de forma diferente em cada canal. Configura-se o cliente para mudar para um determinado canal no momento exacto em que o *beacon* vai chegar. Cria-se então um horário estagnado de períodos de *beacons* para todos os canais. Um cliente pode então basear-se neste horário para ouvir passivamente todos os canais e descobrir APs.

O SyncScan adiciona muitas complexidades, especialmente devido à sincronização de todos os elementos da rede. A forma mais fácil de sincronizar todos os elementos da rede é usar o NTP (*Network Time Protocol*) pela internet, já que fornece um sistema de sincronização de todos os APs pela referência de tempo absoluta. Contudo a sincronização tem riscos, vários APs no mesmo canal tentam enviar o *beacon* no mesmo momento criando interferência. Para resolver este problema a geração do *beacon* pode ser variada aleatoriamente dentro de uma janela temporal (por exemplo 3ms). Um cliente espera durante esse período e ouve os vários *beacons* no mesmo canal. Este processo tem um custo escondido: enquanto remove a carga de rede ocasional imposta por um processo de perscrutação comum, é adicionada carga de rede regular. Enquanto um cliente escuta outros canais não pode trocar pacotes com o seu AP actual. O cliente poderá portanto perder pacotes enquanto sonda outros canais. De seguida apresenta-se a equação que permite calcular o atraso de escuta de cada canal.

$$SyncScanDelay = 2.SwitchTime + WaitTime$$

Este atraso consiste portanto no tempo necessário para mudar de canal, espera durante a janela de tempo pela recepção de *beacons* nesse canal, e por fim o tempo gasto para voltar ao canal de comunicação actual.

Os benefícios principais da abordagem são a redução da latência de *handoff* e conseguir reunir conhecimento da vizinhança para fazer escolhas inteligentes. É ainda sugerido (mas não explorado)

que seja adicionado um dispositivo de localização espacial ao MN para que este, com informação de SNR e RSSI recebida dos APs consiga triangular a localização geográfica de cada AP.

3.3.4 - Eliminating *Handoff* Latencies in 802.11 WLAN Using Multiple Radio [20]

O objectivo principal desta abordagem é eliminar a latência de *handoff* explorando o potencial de rádios múltiplos em dispositivos WLAN. A solução é completamente implementada ao nível do cliente e não requer alterações na arquitectura do AP nem necessita de conhecimento quanto à topologia da rede. Para além da interface principal usada para a comunicação de dados, o cliente usa a segunda interface rádio para sondar oportunistamente o meio, pré-associa-se com APs e eventualmente transferir as comunicações actuais para esses APs tornando o processo de transição de AP transparente para o utilizador.

O uso de duas interfaces em simultâneo num só dispositivo conduz a perdas significativas devido a interferências entre os rádios. Isto ocorre mesmo quando as interfaces se encontram configuradas em canais diferentes e acontece devido à proximidade física dos radios. Assim introduz-se a noção de interface primário para comunicação normal 802.11, enquanto a outra interface, a secundária, é usada para facilitar o *handoff* rápido do tipo *make-before-break* sempre que necessário. Evita-se portanto a comunicação simultânia extendida de ambas as NICs, funcionando a segunda interface apenas em intervalos reduzidos.

Supondo que a inteface primária se encontra associada ao AP1 e é usada para comunicação, a secundária está disponível para desempenhar outras funções. Numa aproximação ingénua a inteface secundária pode executar a fase deperscrutação enquanto a primeira interface comunica normalmente. Assim que a segunda interface encontra um AP apropriado para associação do cliente, a interface primária poderia iniciar o processo de *handoff*. Este *handoff* optimizado poderá então ser executado em menos de 5ms, não contando com o tempo de mudança de canal da interface primária, que demora cerca de 4ms dependendo da interface. Apesar desta não ser a solução óptima e perfeita, a latência de *handoff* é diminuída significamente em contraposição aos tempos de latência actuais.

Numa abordagem mais agressiva é possível eliminar a latência de *handoff* por completo se a interface secundária se associar ao novo AP2 enquanto a interface primária transfere dados através do AP1. Assim que a segunda interface se associa ao novo AP o processo é finalizado, e os papéis das interfaces é invertido entre elas. O funcionamento é então descrito da seguinte forma:

1. *Operação normal:* a comunicação é executada normalmente usando a interface primária associada ao AP1, enquanto a interface secundária executa outras funções, incluindo perscrutação dos canais.

2. *Re-associações:* Se for determinado que a mudança de AP trará benefício, a interface secundária inicia a associação ao novo AP enquanto a interface primária continua em comunicação.
3. *Mudança de interface:* Assim que a segunda interface se encontrar associada ao AP2, todo o tráfego do cliente passa a ser enviado por esta interface. A interface primária torna-se invisível mas continua em funcionamento durante um determinado período para receber pacotes enviados tardiamente pelo anterior AP1.
4. *Conclusão:* As interfaces trocam de papel.

Esta abordagem elimina potencialmente toda a componente de latência do processo de *handoff*. Contudo, a performance em casos específicos deste esquema de *handoff* pode ainda ser negativamente influenciado devido a pacotes perdidos. Pacotes em filas de espera na interface primária serão perdidos se o AP1 tomar conhecimento da associação do MN ao AP2 não aceitando mais pacotes deste utilizador. Isto terá lugar se o canal da interface primária estiver mais congestionado que o canal da interface secundária. Para evitar alterações ao nível do AP ou de qualquer outra estrutura na rede que não o MN, é necessário que ambas as interfaces possuam os mesmos endereços MAC e IP. Desta forma, do ponto de vista da infraestrutura, o *handoff* MultiScan aparenta a existência de um MN com apenas uma NIC que se associa a um novo AP sem latência de *handoff*.

Em contrapartida existem custos de hardware adicionais e maior débito energético, o que não é de todo desejável em comunicações com mobilidade.

3.3.5 - Practical Schemes for Smooth MAC Layer *Handoff* in 802.11 Wireless Networks [25]

Este esquema apresenta duas soluções de descoberta que fundamentalmente ajustam o threshold que controla o disparo do processo de descoberta. Propõe o esquema denominado *Smooth Handoff,* onde a fase de perscrutação dos canais é dividida em múltiplas subfases onde os 11 canais são divididos em grupos. Estes grupos são visitados sequencialmente durante o funcionamento normal da comunicação de dados. O cliente pode usar o intervalo entre duas subfases consecutivas para enviar tramas de dados. Quando todos os canais forem sondados o cliente terá informação sobre todos os APs existentes nos 11 canais (considerando 802.11b/g). Esta informação serve então para fazer parte das políticas de decisão para escolher qual o AP mais apropriado para servir como destino de *handoff*.

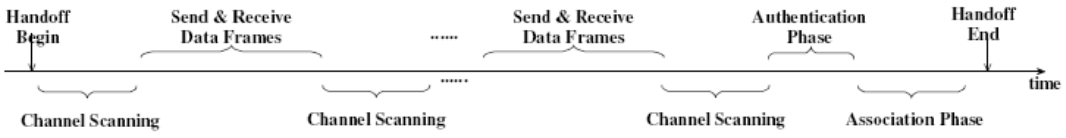

Figura 3.2: Esquema temporal representativo da abordagem do Smooth *Handoff*.

Obviamente, isto pode reduzir o atraso dos pacotes e o jitter durante a fase de perscrutação, que é importante para aplicações críticas como VoIP. A fase de perscrutação deve ser despoletada mais cedo que o esquema de *handoff* implementado actualmente. Para tal usa-se um threshold superior para disparar a fase de descoberta mais cedo de forma a garantir que a desciberta do meio é levada a cabo antes da perda de ligação eminente.

Usando um threshold elevado (*Thres*) para disparar a fase de descoberta poderão verificar-se operações de scan desnecessárias quando o cliente se desloca para uma área geográfica onde o RSSI dos APs decai abaixo desse Thres como consequência de factores externos de atenuação e interferência. Para resolver este inconveniente é implementado um algoritmo adaptativo para alterar dinamicamente o threshold que dispara a fase de perscrutação. Inicialmente, o cliente usa um threshold elevado o que o faz despoletar a fase de descoberta quando o RSSI do AP actual decai abaixo do Thres. Depois desta fase o cliente descobre que o RSSI dos APs na vizinhança é inferior ao Thres. Este valor é então ajustado para um nível inferior (deve ser reduzido α).

Quando um cliente encontra um AP com boa qualidade de sinal o valor do threshold de referência é aumentado β, na expectativa de encontrar um AP com melhor qualidade de sinal. O threshold deve ser limitado superior e inferiormente, $Thresh_{Max}$ e $Thres_{min}$. Isto para garantir que não ocorrem probings excessivos ou evitar que não seja feito nenhum scan e a ligação seja perdida por não ter sido efectuado o *handoff* a tempo.

Este esquema é então alargado a uma versão mais avançada chamada *Greedy Smooth Handoff* onde, adicionalmente, é reduzido o número de canais a ser monitorizado durante a fase de descoberta. Quando o cliente, durante a perscrutação, encontra um AP que reúne as condições necessárias, o processo de scan pára e é logo efectuada a reassociação a este AP. Se não é encontrado um AP adequado no grupo actualmente a ser sondado, o cliente passa ao scan do grupo de canais seguinte. O arranjo dos canais em grupos não segue nenhum critério, pelo que deve haver informação a priori, e os canais com maior probabilidade de conterem APs devem ser agrupados no primeiro grupo a ser sondado.

Para fazer este agrupamento dos canais em grupos, o gestor é obrigado a conhecer toda a rede e ter informação de quais os canais onde os APs estão a funcionar, tal pode ser complexo pois, durante o setup dos APs, estes podem ter a liberdade de escolha de qual o canal mais apropriado para funcionar com o objectivo de melhorar a comunicação global da rede.

3.3.6 - Location-based Fast *Handoff* for 802.11 Networks [26]

Um MN pode derivar quais os APs para os quais é mais provável de ser feito o *handoff* baseando-se na sua localização geográfica actual e na informação da topologia dos APs. Através de um dispositivo de localização espacial (ex: GPS, rede de sensores ou outra técnica), o MN pode obter a sua

localização espacial e em função do movimento e determinar a sua localização relativa quanto a uma certa topologia de APs para efectuar decisões de mobilidade e escolher o destino do *handoff*.

A informação sobre o posicionamento espacial dos APs poderá ser fornecida por um servidor, que também fornece parâmetros para re-associação a cada AP. A diferença entre duas posições medidas consecutivamente fornece um vector de direcção orientado para um grupo **δ** de APs para os quais o MN se irá provavelmente aproximar. O MN escolhe um conjunto **δ** de APs vizinhos para os quais podem ser enviadas mensagens de pré-autenticação ou informação de segurança.

Um servidor de localização organiza a topologia da rede numa tabela de entradas, cada uma para um AP desse domínio, com a forma <parametros, coordenadas (x,y), neighbors>. O primeiro campo inclui informação predefinida em *beacons* para estabelecer reassociações. Os outros campos informam as coordenadas espaciais do AP e os visinhos imediatos. Estas tabelas são introduzidas no servidor de localização manualmente ou automaticamente. O servidor de localização é colocado a par com o servidor de autenticação do domínio pelo que o servidor de localização pode fazer pre-autenticação, distribuição proactiva de chaves ou transferência de contexto para benefício do MN.

Para reduzir a carga da descoberta de posicionamento, o processo é apenas iniciado quando o RSSI desce abaixo de um determinado threshold. Em contrapartida é introduzida mais carga de dados na rede para informar o MN das localizações. São necessárias demasiadas modificações ao nível do cliente. É necessário realizar computação acrescentada e novos dispositivos de localização que consomem bateria já curta de um dispositivo móvel. APs têm de comunicar entre si ou com um novo dispositivo adicionado á rede (o servidor de localização). Propõe o uso de IAPP para efectuar essa comunicação, que é um protocolo que se encontra actualmente a cair em desuso.

Não apresenta a solução que deve ser implementada para que o cliente possa obter informação sobre a sua localização física apresentando apenas sugestões para tal. Não define técnicas para gestão do threshold que dispara a descoberta e cálculo da posição física por parte do cliente. Não apresenta qualquer solução quanto à gestão da segurança na rede, apenas sugere uma transferência de contexto que pode apresentar-se uma actividade complexa.

3.3.7 - Advanced Mechanism for Delay Sensitive Applications in IEEE 802.11 WLAN [27]

Esta abordagem esforça-se em encontrar um threshold óptimo de disparo da fase de descoberta com base na medição do RSSI. A solução é nomeada FSFR-HO (*First Satisfaction First Reservation Handoff*). O cliente sonda os canais até encontrar o primeiro AP com um RSSI acima do threshold definido chamado *Satisfactory Threshold*. Depois o cliente passa directamente para a fase de reautenticação em vez de sondar todos os canais à procura do AP com melhor RSSI. O valor para o S*atisfactory Threshold* pode ser o mesmo que o *Handoff Threshold* usado para disparar o *handoff*, isto porque o

cliente estava satisfeito enquanto o RSSI do antigo AP estava acima deste valor e também porque assim mantém-se a simplicidade da implementação já que o valor já está definido na camada MAC.
Se o RSSI é apenas um pouco superior que o threshold definido, o novo AP é considerado como sendo adequando. Contudo o decaimento da qualidade do sinal pode ser rápido, deixando o cliente novamente insatisfeito com o serviço, ocorrendo de novo mudança de AP. Estas mudanças de APs podem causar o efeito de "ping-pong" entre APs ou mudanças desnecessárias que poderiam ser evitadas caso o AP escolhido fosse o ideal.

3.3.8 - Pre-Scanning and Dynamic Caching for Fast *Handoff* at Mac Layer in IEEE 802.11 WLAN [28]

Este trabalho propõe o uso de um mecanismo de pre-scanning com uma máscara selectiva de canais para a fase de descoberta activa feita em segundo plano e um mecanismo de caching dinâmico. Neste esquema de *handoff* cada canal é sondado antes da fase de *handoff*. Esta abordagem tenta eliminar o atraso de perscrutação usando um algoritmo de perscrutação proactiva e demora apenas alguns milisegundos usando um sistema de caching dinâmico. Estes processos têm lugar no cliente e não são necessárias alterações na rede ou no standard 802.11, apenas são necessárias algumas alterações no driver da NIC do cliente.

Este trabalho aborda o problema baseando-se nas seguintes premissas.

- Se o número de canais a serem sondados é reduzido então reduz-se o atraso.

- Se o cliente conhecer todos os APs a serem sondados de experiências de fases de descoberta anteriores pode parar a espera de respostas no respectivo canal antes que o *MinChannelTime* expire.

- Se o cliente já conhecer o AP para onde pretende transferir a associação, então não será necessário fazer a descoberta do meio.

Perscrutação Selectiva: Apenas os canais seleccionados são sondados. Quando o cliente é activado faz um scan complecto e preenche a máscara. Este processo é então iniciado novamente, em segundo plano, quando a força de sinal desce abaixo de um threshold definido (mas superior ao threshold do *handoff*). Um exemplo de uma máscarade canais é apresentado de seguida.

Figura 3.3: Exemplo de uma máscara de canais usada neste mecanismo.

Cada bit '1' na máscara indica que existem APs no respectivo canal.

Caching: Usando uma cache, a descoberta do meio pode ser evitada e passa-se directamente para a autenticação e associação. Este mecanismo cria uma tabela que catalóga dois APs adjecentes para cada AP. Quando acionado o *handoff* o cliente consulta esta tabela. Se for encontrado um AP, o cliente envia-lhe um pedido de associação. Quando o cliente não consegue a ligação ao primeiro AP da lista então tenta a segunda hipótese, se esta também falhar inicia-se a perscrutação selectiva. Usando a cache a latência do *handoff* diminúi para poucos milissegundos. Contudo o caching não garante uma latência de *handoff* constante porque as entradas da cache resultam do historial do cliente em movimento e não de resultados actuais (podem ocorrer mudanças na rede). Se alterações ocorrem a latência sobe acima de 100ms persistindo o problema.

Apenas baseia o disparo do *handoff* na qualidade de sinal. A máscara de perscrutação selectiva é apenas preenchida no início da comunicação do MN na rede, ou seja, como o MN se desloca por toda a rede, novos APs podem demonstrar-se melhores opções, mas não são sondados pelo MN. A cache pode ficar desactualizada se o MN se desloca para locais afastados daquele onde a cache foi elaborada.

3.3.9 - Selective Channel Scanning for Fast *Handoff* in Wireless LAN using Neighbor Graph [29]

O objectivo desta abordagem é reduzir a latência do processo de perscrutação pelo uso de NGs. Nesta abordagem é levado a cabo a perscrutação selectiva de canais em unicast, no qual o cliente apenas sonda APs específicos seleccionados pelos NGs. O cliente envia uma trama de *Probe Request* em unicast para o AP num determinado canal referenciado pelo NG e após receber a resposta muda logo para o canal seguinte sem esperar o tempo pré-definido que obriga o cliente a manter-se no canal para receber esta resposta (*MinChannelTime*) nem o tempo para receber eventuais respostas de outros APs (*MaxChannelTime*). Desta forma o atraso por canal na fase de perscrutação é reduzido. O tempo de descoberta do meio pode então ser traduzido pela seguinte expressão:

$$t = N' \times rtt + \alpha$$

onde **N'** é o número de potenciais APs apontados pelo NG para serem sondados, **rtt** é o tempo de envio e recepção do probe unicast (*round trip time*) e α é o tempo de processamento da mensagem.

Conta que a fase de perscrutação seja reactiva, isto é, tem início quando a qualidade de sinal decai abaixo de um threshold que marca o início da fase de *handoff*. A preparação do *handoff* consiste na manutenção dos NGs, e não considera qualquer outra métrica de avaliação para as decisões de *handoff*.

A forma como os NGs são construídos e geridos não é abordada. A manutenção e disseminação da informação do NG representam encargos de gestão potencialmente elevados para aqueles que fazem

parte de uma rede 802.11 alargada. São necessárias alterações ao nível dos clientes e APs, não considerando a necessidade de uma entidade que armazene e disponibilize informação do NG.

3.3.10 - Improving Latency of 802.11 *Handoff*s using Neighbor Graphs [30]

Esta abordagem faz uso de um método de descoberta por neighbor graphs e non-overlap graphs. Este método reduz o número total de canais sondados e o tempo de espera em cada canal. Baseia-se no princípio que o maior contribuinte para a latência do processo de *handoff* consiste em identificar o novo melhor AP disponível e associar-se a esse AP. A fase de perscrutação é afectada por dois parametros, pelo *probe count* (número de canais sondados) e *probe-wait time* (tempo gasto em cada canal á espera das tramas *probe response* dos APs).

São portanto propostas duas abordagens ao problema de *handoff* na camada 2:

-Algoritmo NG: pelo uso de neighbor graphs.

-Algoritmo NG-pruning: pelo uso de non-overlap graphs.

Algoritmo de NG: Tendo conhecimento do NG o MN pode sondar apenas os canais não vazios (com APs a operar), resultando num *probing count* reduzido por exclusão dos canais de desperdício (retiram-se os canais vazios), obtendo-se um *minimum-channel probing*. Este algoritmo também diminui o *probe-wait time* saltando de canal logo assim que o AP do canal actual responde, resultando num tempo de espera chamado *optimal-wait probing*, assim como é proposto pelo trabalho [29]. A geração dos NG é efectuada por meio dos seguintes métodos: *edge-addition*, em que, quando um cliente muda do APi para o APj e <APi, APj> $\notin$NG, adiciona o par ao NG. O segundo método denomina-se *edge-delection*, que consiste na remoção do par <APi, APj> do NG quando nenhum cliente efectua *handoff* do APi para o APj durante um intervalo **T**. Este período **T** é um timeout escolhido de forma que **T**≥tempo médio de *handoff* entre todas as associações de APs.

NG-Pruning Algorithm: Usa non-overlap graphs para ganhar um suplemento de performance na perscrutação. O cliente exclui (prune) todos os APs que não se sobrepõem a APs acessíveis.

Estas políticas de geração de NGs implicam que estes não possam ser usados nos primeiros tempos de vida da rede, contudo, com o decorrer do tempo, a performance do processo melhora.

A preparação do handover consiste na geração e manutenção dos NGs e non-overlap NGs. As decisões de *handoff* levam em conta o resultado da sessão de perscrutação. Para redução da latência de *handoff* apenas é melhorada a fase de perscrutação. Não existe nenhum NG à partida começando este por ser uma tabela vazia. Nos primeiros processos de *handoff*, estes são efectuados sem qualquer informação fornecida pela nova abordagem, com pesquisa do meio por perscrutação activa disparado por um threshold tal como na forma tradicional.

3.3.11 - Context Caching using Neighbor Graphs for Fast *Handoff* in a Wireless Network [31]

Esta abordagem implementa uma estrutura de dados , o NG, que dinamicamente captura a topologia de mobilidade de uma rede como meio de pré-disponibilizar o contexto do cliente assegurando que este contexto se encontra sempre disponível no AP seguinte. Esta transferência de contexto é efectuada proactivamente e não reactivamente, i. e., antes da perda de comunicação e não disparada por este evento. O problema das abordagens proactivas reside na dificuldade de determinar quando deve ser despoletada esta transferência de contexto e determinar qual o conjunto de potenciais APs a serem notificados com este contexto. O NG disponibiliza o conjunto de APs na vizinhança com cobertura para os quais o cliente vai provavelmente migrar. A cache está implementada em cada AP para armazenar o contexto dos clientes com probabilidade de realizar *handoff* para a sua BSS, ou quando o MN está prestes a fazê-lo. Quando o MN notifica a intensão de handover já terá sido feita a transferência de contexto para todos os APs que se encontram catalogados como sendo os vizinhos do AP actual mais prováveis de serem o alvo do *handoff*.

A preparação do *handoff* consiste na criação e manutenção dos NGs, que não é clarificada neste documento. Para redução do tempo de *handoff* são utilizados os NGs que reduzem a latência da fase de perscrutação, e implementa-se o caching nos APs destino contendo a informação necessária para retomar a comunicação normal do MN. Para decisões de *handoff* é utilizada a métrica da qualidade de sinal, o AP que apresentar o melhor RSSI será o seleccionado de entre os outros potenciais APs. As principais contrapartidas consistem no uso de IAPP para transferir dados entre os APs e na elevada carga incumbida aos APs que devem armazenar informação de inúmeros clientes. O contexto de cada MN é difundido por todos os APs na vizinhança descritos no NG, o que pode revelar-se obsoleto e ainda mais sobrecarregar a cache limitada de cada AP.

3.3.12 - A Selective Neighbor Caching (SNC) Scheme for Fast *Handoff* in IEEE 802.11 Wireless Networks [32]

Este trabalho segue a ideia principal apresentada pela abordagem [31]. Um AP propaga proactivamente o contexto do cliente aos APs vizinhos. Contudo, levando em conta a probabilidade de *handoff*, envia o contexto do cliente apenas aos APs que têm uma probabilidade de *handoff* igual ou superior a um valor predefinido. Os APs com probabilidade inferior a este valor não receberão o contexto de comunicação do cliente. Isto poderá representar um problema desta implementação, uma vez que o cliente pode transitar para um AP onde não existe ainda o contexto necessário, ocorrendo um *cache miss* resultando numa latência de *handoff* superior. No SNC, cada AP atribui um peso a cada um dos APs na sua vizinhança com base na probabilidade de *handoff* de si para estes APs. Mediante este peso, o AP propaga o contexto do cliente apenas aos APs com peso igual ou superior a um valor predefinido (**δ**). O peso de cada vizinho pode ser obtido através do processo de construção do NG e é armazenado na tabela denotado por *W=[wi(j)]*, onde **wi(j)** é a probabilidade de transição

do **api** para o **apj**. Seja **Ci(j)** o número de eventos de *handoff* do **api** para o **apj** durante o tempo de monitorização, então:

$$w_i(j) = \frac{C_i(j)}{\sum_{all \quad neighbor \quad k} C_i(k)}$$

E é feita a propagação do contexto do MN para os APs que seguem a condição **wi(j) ≥ δ**.

Mais uma vez, não é revelado como é criado o NG nem abordadas políticas de ataque à criação ou gestão de chaves para protocolos de segurança.

3.4 - Optimização do Handoff com Requisitos de Segurança

O tratamento da informação referente aos mecanismos de segurança é um mecanismo moroso, o que afecta directamente a experiência de mobilidade do MN. Na eventualidade de um *handoff* de um MN a implementar o protocolo de segurança IEEE 802.11X, é experimentado atraso de estabelecimento da sessão com o novo AP na ordem dos segundos, como apresentado na tabela 3.1 deste capítulo. Assim surge a necessidade de mecanismos alternativos que permitam a redução desta componente de latência que afecta gravemente o *handoff*.

Como explicado no capítulo 2, o protocolo 802.1X organiza uma hierarquia de chaves que possibilitam a segurança da comunicação. Para tratar da geração e distribuição da hierarquia de chaves na situação de *handoff* existem três abordagens possíveis:

1. Transferência proactiva do contexto de segurança entre APs
2. Recriação rápida de novos contextos de segurança em novos APs.
3. Abordagens híbridas.

A gestão de chaves é o aspecto principal a ser manipulado na implementação de um protocolo de segurança. Resta apenas decidir como e quando realizar os mecanismos necessários para que, no final do processo, tenham sido criada e distribuída a hierarquia de chaves de forma funcional e eficaz.

Apresentam-se de seguida as abordagens de diversos autores estudadas na execução deste trabalho.

3.4.1 - Roaming Key based Fast Handover in WLANs [33]

Para além de uma solução de *handoff* dentro do mesmo domínio (intra-domínio) também aborda a questão do *handoff* entre diferentes domínios (inter-domínio). Esta solução faz uso dos padrões IEEE 802.11F e IEEE 802.11i. As chaves em uso são tal qual definidas no 802.11i, nomeadamente uma *Master Key* (MK) partilhada pelo cliente e pelo servidor AAA, uma PMK no AP e no MN derivada da MK enviada pelo servidor AAA, uma PTK e uma GTK negociadas entre o AP e o cliente.

Apresenta uma solução que faz uso de uma *Roaming Key* (RK) para proporcionar autenticação mútua rápida quando ocorre um *handoff*. O método de obtenção desta chave não é proposta, apresentada nem discutida pelos autores deste trabalho. Esta chave pode ser usada para cifra no *handoff* de inter-domínio enquanto a PMK e PTK não forem criadas. A comunicação irá então continuar até que a chave RK expire. Esta RK beneficia o *handoff*, usada em conjunto com informação de contexto (*Context Information,* CI), transferida do AP actual para o AP alvo, e informação de segurança (*Security Information*, SI), enviada ao domínio destino (ou rede alvo) pelo domínio de serviço (ou rede actual).

O CI incluí o tipo de cifra usado pelo cliente, PMK e tempo de expiração (Time-Out, TO) TO_{PMK}, RK e o TO_{RK}, PTK e TO_{PTK} se esta não é derivada a cada vez que ocorre um *handoff*, TO_{CI}, timestamp, endereço MAC do AP e o ID do cliente (ID temporário criado durante o primeiro login). O SI distingue-se do CI por ser informação de segurança da rede de destino. O AP actual envia o SI ao cliente quando existe um *handoff* para um domínio diferente. Contém informação como o ID da rede de destino (ESSID e BSSID), o *Care-of-Address* (CoA) caso esteja a ser usado o MIP, ou o IP novo para o cliente, as chaves e os seus períodos de TO e os algoritmos de cifra que podem ser usados na rede destino.

A SI consiste em duas chaves, a RK e a chave de cifra, e a nova PTK. O AP no domínio destino e o cliente partilham estas chaves. A distribuição da SI pode ser efectuada das seguintes formas: todos os domínios já possuem as SIs uns dos outros ou então são disponibilizados por parte do servidor AAA em cada domínio e é requisitado por pedido do AP actualmente a servir o cliente.

De seguida é explicado o percurso do protocolo num ambiente intra-domínio, que é o relevante para o nosso trabalho.

***Handoff* Intra-Domínio:** O AP1 liga-se ao servidor AAA e cria a associação de segurança sendo estabelecida a ligação de segurança com os outros APs da rede sugeridos pelo NG. O cliente negocia os métodos de cifra e autenticação e liga-se ao servidor AAA. Nesta interacção é criada a PMK que é enviada ao AP1. O cliente e o AP1 derivam a PTK e a RK partindo da PMK. Quando o MN começa a comunicação através do AP1 envia-lhe o seu CI para ser difundido pelos outros APs através da associação de segurança criada inicialmente. No momento de *handoff* o cliente notifica o AP1 da sua intenção de mudar de AP. Finalmente, o cliente e o AP destino efectuam autenticação mútua usando a RK pré-distribuída e retoma a comunicação normal e o novo AP informa o servidor AAA sobre o *handoff*. Todas as chaves devem ter um prazo de expiração. A renovação da RK deve apenas ser

pedida durante ou após um *handoff*. A PTK pode ser criada em cada AP após o *handoff* enquanto a comunicação é efectuada com o uso da RK. Se o novo AP de destino não constar do NG não irá ter a RK para o cliente pelo que será necessária uma autenticação mútua genérica. Apesar de isto aumentar o atraso do *handoff* garante que o cliente não se irá ligar a um AP falso.

3.4.2 - Fast Pre-Authentication Based on Proactive Key Distribution for 802.11 Infrastructure Networks [34]

Este trabalho propõe dois métodos de reautenticação rápida baseados num mecanismo preditivo de autenticação definido pelo grupo de segurança IEEE 802.11i.

3.4.2.1 - Distribuição Proactiva de chaves (PKD)

Este método efectua a distribuição proactiva de chaves entre o cliente e os APs. Desta forma são criadas chaves de autenticação antes das reassociações. No momento do *handoff* a troca de chaves entre o cliente e o AP reduz-se ao 4WHP e ao Group Key Handshake. Este processo baseia-se num servidor AAA responsável por gerir um NG para todos os APs da rede. Após a primeira autenticação mútua 802.1x entre o cliente e o servidor AAA, o AP envia ao servidor AAA uma mensagem *Accounting-Request (Start)*. Com isto o servidor AAA informa a vizinhança de APs, através de um *Notify-Request*, que daquele AP específico poderá ocorrer um *handoff* de um determinado cliente. A esta altura, cada AP responde ao servidor AAA com um *Notify-Response* que inicia o processo de criação de uma nova chave PMKn criada a partir de uma função PRF aplicada sobre a PMK inicial ou do *handoff* anterior e dos endereços MAC do cliente e AP destino do *handoff*. Feito isto o servidor AAA envia as chaves a cada AP usando uma mensagem *ACCESS-ACCEPT*. No momento de *handoff* para um novo AP, o cliente obtém o endereço MAC desse AP e calcula a PMKn. Esta será igual à chave PMKn já enviada pelo servidor AAA a cada um dos APs e diferente para cada AP (pois possuem diferentes endereços MAC). Depois disto apenas será necessário efectuar um 4WHP e um Group Key Handshake para validar as chaves e iniciar a comunicação.

Este método por si só apenas elimina a comunicação com o Servidor de Autenticação pelo que, com vista a reduzir a latência suficientemente de forma a permitir a implementação deste protocolo com aplicações de tempo real, este método é concatenado com dois mecanismos seguidamente expostos:

3.4.2.2 - Mecanismos de complemento ao PKD

• PKD com caching IAPP: Para além da pré-distribuição de chaves PKD, usa o processo de transferência de contexto do IAPP para distribuir as chaves PTK. O cliente usa estas chaves para se reassociar temporariamente ao novo AP sem necessidade de efectuar o 4WHP. Apenas é necessário efectuar um simples Group Key Handshake. As chaves PTK pré-distribuídas são calculadas pelo AP actual e enviadas aos outros APs. Estas chaves são calculadas aplicando uma função PRF sobre a PMK

actual, a PTK inicial e os endereços MAC do cliente e de cada AP. O cliente poderá calcular a PTK assim que souber o endereço MAC do AP destino do *handoff*. Esta chave temporária irá servir para garantir comunicação contínua do cliente com o novo AP. Enquanto um temporizador instituído no início deste processo não terminar, o cliente pode continuar a comunicar na rede ao mesmo tempo que efectua o 4WHP original para obter a PTK definitiva. Este método tem a desvantagem de usar IAPP que, como já foi dito, caiu em desuso.

• PKD com 4WHP antecipado: Este método tem a finalidade de eliminar a fase do 4WHP da fase de reautenticação e o abandono do protocolo IAPP resumindo a reautenticação ao Group Key Handshake (2 mensagens). O servidor AAA envia ao cliente uma lista de APs na vizinhança que responderam ao Notify-Request na troca PKD. Durante a associação ou reassociação, o AP actual sinaliza o servidor AAA de forma a iniciar a distribuição de chaves proactiva. Todos os APs da vizinhança receberão as chaves PMK relativas ao cliente. O cliente irá conduzir uma pré-autenticação através do sistema de distribuição via AP actual com os APs da lista de vizinhos da rede. O cliente efectua o 4WHP com estes APs com a PMKn calculada graças ao conhecimento dos endereços MAC dos APs disponibilizados pela lista de APs vizinhos fornecida pelo servidor AAA. Quando o cliente se desloca em direcção a um destes APs vizinhos podem ocorrer duas situações:

- O cliente já calculou a PTK na pré-autenticação e apenas efectua o Group Key Handshake para completar a autenticação.

-O cliente ainda não ainda não finalizou a pré-autenticação, logo terá de efectuar o 4WHP e o Group Key Handshake correspondente ao processo completo do método PKD.

É necessário o conhecimento de toda a rede para implementação de NGs no servidor AAA. Isto torna-se impraticável para redes de largas dimensões e implica modificações complexas nesse servidor.

3.4.3 - CAPWAP Handover Protocol [13]

O processo de *handoff* é gerido por um controlador de acesso (*Access Controller, AC*) ligado aos APs de um domínio. Este AC efectua a transferência proactiva de contexto entre APs, ou seja, antes que o *handoff* se realize, com a finalidade de eliminar o atraso associado a esta fase do processo de *handoff*. O CAPWAP introduz um novo conceito de APs construídos com a arquitectura Split MAC aos quais se dá o nome de *Wireless Termination Point* (WTP). Estes WTPs na arquitectura Split MAC implementam os serviços MAC sensíveis ao atraso (geração de *beacons*, resposta a *probe requests*, etc.) e direccionam por túnel todos os dados e algumas tramas de gestão (associação) ao AC para processamento centralizado. No CAPWAP é necessário um novo protocolo de *handoff* de forma a gerir centralizadamente o processo pelo AC e incorporar os WTPs Split MAC.

O CAPWAPHP (*CAPWAP Handover Protocol*) foi desenhado para ser o protocolo de comunicação inter-AP para o CAPWAP. Permite também, além da transferência de contexto da comunicação e

associação actual do MN, a transferência de contexto AAA entre WTPs e assegura que as tabelas de switching são actualizadas no DS. O CAPWAP detém um NG centralizadamente. Este NG pode ser criado estaticamente e introduzido na base de dados do AC ou gerado dinamicamente segundo qualquer outro processo. Quanto ao contexto de segurança, contando que foi gerada uma PMK na primeira autenticação do MN na rede, pode ser reutilizada a PMK que é transferida entre WTPs que, processada com o MAC do novo WTP, resulta numa nova PMK. Esta derivação irá suceder-se à medida que o MN se desloca pela rede e altera a sua associação entre WTPs. Chaves PMK derivadas desta forma tornam-se parte do contexto AAA.

Na primeira associação do MN com o primeiro WTP é encaminhado o pedido ao AC, onde é processado. O AC responde ao WTP incluindo os valores e informações relevantes para o *handoff* e o WTP, por sua vez, responde ao MN atendendo ao pedido de associação. É então executada a autenticação 802.11i entre o MN e o servidor AAA de onde resulta uma PMK. De seguida o AC difunde o contexto de segurança do MN para todos os WTPs na rede que, por mecanismos de caching, armazenam o contexto durante a sessão activa para possíveis futuras transições do MN específico. Na eventualidade do *handoff*, o contexto já se encontra nos WTPs que poderão vir a ser o alvo do *handoff*. O novo WTP faz o pedido de transferência do contexto de associação ao AC que trata desta troca com o antigo WTP. Após feita esta transferência, e contando que o contexto de segurança já teria sido difundido pela rede, é levado a cabo o 4WHP entre o MN e o novo WTP para verificação e derivação de chaves. Na eventualidade de o contexto já ter sido previamente transferido, não é feito qualquer pedido ao AC e apenas é feita a verificação de segurança com o 4WHP. Isto oferece a possibilidade de adaptar o protocolo CAPWAPHP a um cenário com ou sem caching da informação de associação e comunicação. Se todos os WTPs são Split MAC, não é necessário o caching do contexto de segurança e este é mantido apenas no AC.

3.4.4 - Personal AP Protocol for Mobility Management in IEEE 802.11 Systems [35]

No sistema de suporte de mobilidade “Personal AP”, o contexto de mobilidade de cada MN é definido pela informação de estado relevante respeitante ao AP actual, incluindo os estados de associação da camada MAC no AP. Quando o MN muda de AP o seu contexto segue-o de uma ligação física para outra, criando assim um AP fantasma que persegue o MN, evitando a latência de reassociação com novos APs. A latência do *handoff* é então eliminada evitando o processo de reassociação moroso característico dos sistemas 802.11.

3.4.4.1 - Decisões de *handoff*: O Personal AP baseia as decisões de *handoff* em duas características da rede:

- Qualidade de sinal nas ligações estabelecidas no momento. Geralmente o RSSI é o mais indicado para obter esta informação. Deve ser tomada a ligação com melhor RSSI.

• Características do tráfego das aplicações a correr no MN. Diferentes ligações têm diferentes taxas de transmissão, envios em blocos volumosos e diferentes durações. Assim, MNs com aplicações exigentes devem implementar os mecanismos do Personal AP, enquanto MNs a correr aplicações pouco exigentes podem efectuar o *handoff* de qualquer outra forma pois a latência característica não representa um impacto significativo nessas aplicações.

O AP está encarregue de reunir informação de RSSI medido nas tramas dos MNs para fornecer informação das ligações ao controlador de acesso (*Access Controller*, AC). Por simplicidade, uma média exponencial do RSSI é armazenada no AP. O tipo de dados também é notificado ao AC, sendo este que decide a necessidade de implementação do Personal AP iniciado e assistido pela rede ou implementação de outro processo de *handoff* genérico.

3.4.4.2 - Operações de *handoff*: Quando é tomada a decisão de realizar o *handoff*, o Personal AP transfere a informação de contexto de um AP para outro. Na primeira associação do MN o AP actual informa o AC sobre essa associação e este cria um mapeamento desta ligação específica de cada MN. Quando o RSSI medido na proximidade de um novo AP supera o RSSI medido no AP da associação actual do MN, o AC inicia o *handoff* assistido pela rede transferindo a associação do AP actual para o novo AP, se a melhoria de performance prevista o justificar. O Personal AP implementa a transferência de contexto de forma similar ao IAPP. Assim são usadas mensagens *Watch* para notificar periodicamente a qualidade de sinal ao AC e mensagens *Start* para ordenar ao novo AP que se comporte como o AP a que o MN se encontra associado e requisitar ao AP actual a transferência de contexto para o novo AP. Mensagens de *ACK* e *Done* são usadas para notificar e confirmar recepções e finalizações de mensagens e processos.

Este processo distingue-se do IAPP pois não necessita de um NG nem distribuição de várias cópias do contexto do MN para variados APs.

3.4.4.3 - Transferência do contexto de segurança: Se for implementado o 802.11i, chaves relevantes como a PMK são transferidas com o contexto de cada MN para cada AP sempre que seja necessário ocorrer o *handoff*. O servidor AAA, no final do 802.1X da primeira associação do MN na rede, transfere a PMK para o AC e este responsabiliza-se pela sua difusão a cada AP a que o MN se associe. Em cada nova associação a um novo AP, com o conhecimento da PMK e pelo 4WHP, são derivadas as chaves necessárias.

3.4.4.4 - Comunicação de dados: Com o Personal AP é efectuada a transferência completa do contexto para o novo AP, pelo que este pode comportar-se como se fosse o AP anterior e actuar como se o AP anterior tivesse seguido o MN e nenhum *handoff* tivesse ocorrido. Por este motivo nada precisa ser alterado nas tramas de dados ou na forma como estas são difundidas entre o cliente e o AP.

3.4.5 - A Seamless *Handoff* Mechanism for IEEE 802.11 WLANS Supporting IEEE 802.11i Security Enhancements [36]

Na autenticação 802.1x tal como esta se encontra implementada actualmente, os clientes não podem enviar/receber dados pelo novo AP enquanto a autenticação não tenha sido bem sucedida. Este trabalho propõe que um cliente possa comunicar normalmente através do novo AP, durante um curto período de tempo, enquanto a autenticação 802.1x é feita.

Este trabalho sugere o estabelecimento de um túnel dinâmico como solução para atingir o *handoff* transparente. O estabelecimento de túneis dinâmicos irá permitir que pares de APs comuniquem bidireccionalmente de forma fidedigna. Os clientes poderão usufruir dos serviços proporcionados pelo AP ao qual estavam anteriormente associados, via AP actual, entre os quais existe um túnel dinâmico em funcionamento. Esta comunicação tem lugar durante um tempo limitado enquanto se processam as fases do *handoff* para reassociação ao novo AP.

A comunicação não é propriamente efectuada pelo novo AP, mas é sim encaminhada para o AP antigo através de um túnel estabelecido no backbone da rede. Os túneis permitem aos APs verificar se os APs vizinhos a que se ligam são de confiança. Enquanto o cliente é autenticado com o servidor de autenticação via novo AP, a comunicação continua com o AP anterior para que qualquer trama destinada ao cliente possa ainda ser reencaminhada. Assim é possível obter um *handoff* rápido.

A criação dinâmica de túneis é iniciada pela recepção de uma trama de *reassociation request* no novo AP. O novo AP descobre que o AP ao qual o cliente está associado actualmente se encontra na sua vizinhança e tenta então estabelecer um túnel. Antes de enviar o pedido ao AP anterior, o novo AP verifica se o anterior é de confiança interrogando o servidor de autenticação. Se a verificação for bem sucedida, o novo AP envia uma mensagem *tunlestb-request* ao AP anterior, o que incentiva o AP anterior a verificar a autenticidade deste novo AP com o servidor de autenticação e a construir um túnel com esse AP.

Como a criação de um túnel é despoletada pelo *handoff* de clientes, o primeiro cliente não irá de imediato usufruir deste túnel pois é criado em paralelo com a autenticação 802.1X, não havendo comunicação contínua durante o *handoff*. Contudo, clientes que posteriormente alterem a sua associação entre este par de APs podem usufruir deste túnel e utilizar esta nova abordagem tal como ela foi concebida.

A complexidade dos APs aumenta com a implementação deste método pois, agora, terão que decifrar e validar tramas da camada 2 provenientes do MN e do túnel seguro bem como devem cifrar tramas da camada 2 para que sejam enviadas via sem fios e pelo túnel seguro. Reassociações falsas ou repetidas por atacantes de uma rede podem iniciar a criação de túneis desnecessária.

3.5 - Análise e Avaliação das Abordagens apresentadas

A transferência de contexto entre APs demonstra-se apelativa uma vez que reduz [33], [34] e [13] ou faz mesmo desaparecer [35] a carga computacional necessária para a reassociação.

Algumas arquitecturas propostas gerem os APs como sendo terminais físicos de um "AP central" gerido pela rede, onde o contexto de segurança dos MNs pode migrar para outros APs que se encontrem próximos do MN.

Contudo a transferência de contexto apresenta também diversas desvantagens. A primeira desvantagem é que a rede de APs deve conter algum tipo de gestão na infra-estrutura para tratar da migração segura de MNs entre APs. [33] usa um NG de APs e precisa que sejam estabelecidas associações de segurança entre APs arbitrários. [34] usa informação conferida por um AS e informação sobre a vizinhança de APs. [35] usa um controlador de acesso centralizado que se encarrega de controlar todas as decisões respeitantes ao *handoff*. [13] usa uma arquitectura CAPWAP centralizada que envolve entidades de rede extra e requer instalações para gerir tabelas de switching. Estas infra-estruturas levantam diversos problemas de segurança que são descritos em [37].

A segunda desvantagem é que, segundo o 802.11i, a hierarquia de chaves usada em cada AP para um determinado MN (começando na PMK) deve ser diferente, o que implica que em qualquer caso, o AP e o MN devem levar a cabo o 4WHP após a reassociação [33], [34] e [13]. Em [34], no entanto, é proposta uma alternativa para adiar o 4WHP, reutilizando temporariamente a PTK distribuída pelo AP do qual o MN transferiu a associação. Efectuar o 4WHP é obrigatório se os elementos RSNIE conferidos pelos diferentes APs forem diferentes. Por outro lado, em sistemas em que os MNs não diferenciam quando estão a ser servidos por diferentes APs, tal como ocorre em [35], torna-se bastante complexo, se não impossível, empregar APs com capacidades operacionais diferentes.

A terceira desvantagem é que os APs devem proactivamente transferir os contextos antes que se efective a reassociação, onde podem ser gastos recursos e tempo a tratar de um problema que poderá nem vir a ser verificado. Ainda mais, este esforço pode não ser suficiente, uma vez que em [34], [33] o MN deve correr a autenticação 802.1X completa sempre que se associa a um AP que não se encontre no NG do AP que o serve actualmente.

Por fim, a última desvantagem é permitir que MNs sejam servidos de forma transparente por diferentes APs, tal como em [35], o que complica a gestão da rede de acesso, nomeadamente a gestão de routing e tabelas de tradução de endereços da camada 2.

Quando os contextos de segurança são recriados nos novos APs, para implementar *handoff* rápido torna-se necessário efectuar optimizações para reduzir o atraso imposto por fases pós-reassociação 802.1X. Em [36] é proposta uma arquitectura onde um MN pode comunicar após estar reassociado e antes de correr as fases que faltam do protocolo 802.1X (fases 2 e 3 descritas no capítulo 2). A comunicação é encaminhada por túnel ao AP anterior através de túneis dinâmicos seguros (Dynamic

Secure Tunnels). Estes túneis são criados a pedido durante a reassociação, com a ajuda do AS, e são mantidos posteriormente entre os APs para servir MNs que tenham efectuado transição entre o mesmo par de APs. Contudo, esta abordagem provavelmente complicará o funcionamento dos APs, uma vez que estes deverão decifrar e validar tramas da camada 2 provenientes da interface sem fios e de túneis seguros, e vice-versa. Pedidos de reassociação manipulados por mão de um atacante podem ainda levar à criação de túneis sem utilidade.

Em [38] é proposta uma arquitectura baseada num protocolo de distribuição de chaves segura com três intervenientes, usando um servidor HOKEY local [39] além das entidades 802.1X convencionais, um autenticador e um AS. O servidor HOKEY reduz a duração da segunda fase do 802.1X, substituindo a autenticação EAP completa por uma autenticação ERP local (EAP Reauthentication Protocol [40]) entre o MN e o servidor HOKEY. O servidor HOKEY usa material de chaves derivado do EAP para o ERP, conferido pelo AS local, eliminando assim pedidos repetidos ao AS remoto. Contudo continuam a existir as mesmas fases do 802.1X: fase 2, agora com ERP e fase 3 (4WHP).

Em [41] é proposta uma abordagem híbrida, onde parte do contexto migra entre APs e outra parte é recreada entre o MN e o novo AP. Os APs são dinamicamente organizados em grupos usando NGs e em cada grupo existe uma chave de roaming (*Cluster Roaming Key*, CRK) por cada MN que o visita. Esta CRK é calculada da chave PMK inicial do MN e da lista actual de membros do grupo. Esta CRK é usada pelo MN e AP para derivar a própria PMK local sem troca de mensagens. Usando uma PMK por AP, um MN executa um processo equivalente ao 4WHP usando apenas duas tramas de reassociação autenticadas (pedido e resposta). Contudo, usar apenas duas mensagens requer a geração autónoma de nonces: o MN deve adivinhar o nonce que o AP irá usar para calcular a PTK da PMK. Esta previsão pode falhar obrigando à troca de duas tramas adicionais de sincronismo. Apesar das semelhanças desta abordagem com o nosso protocolo proposto, tal como o uso de protocolos 802.11 para autenticação 802.1X e autenticação das reassociações, esta abordagem tem contudo que usar agrupamentos de APs e fornecer informação sobre esses grupos aos MNs, para que estes possam calcular CRKs, preocupações que não temos na nossa abordagem. Adicionalmente, eles não implementam reassociação rápida quando o MN migra para APs de grupos distintos. Por fim, um AP sob influência de um atacante num determinado grupo pode derivar a PMK de um MN usado por todos os outros APs no mesmo grupo, enquanto na nossa abordagem a PMK usada em cada AP não pode ser usado na derivação da PMK noutros APs.

O standard 802.1r visa reduzir a carga de segurança durante a migração de MNs pelos APs de uma rede. As duas contribuições mais evidentes deste protocolo em termos de segurança são a clarificação de caching de chaves oportunistas em APs e a dispensa do 4WHP que sucede as reassociações no protocolo 802.1X actual. O 802.11r define uma nova hierarquia de chaves baseadas na EMSK e o conceito de domínio de mobilidade. Um grupo de APs constitui um domínio de segurança e com isto ganhando acesso a uma hierarquia de chaves baseadas na EMSK comum para cada MN. Num cenário óptimo, um MN assume que tal hierarquia de chaves já existe no AP destino da transição quando o *handoff* ocorre (caching oportunista de chaves). Se tal não acontece é

experimentada uma latência inesperada enquanto o AP obtém essa informação da entidade que a detém. O 802.11 não especifica como as chaves são distribuídas para os APs do mesmo domínio de segurança e como os APs podem requerer hierarquias de chaves a pedido. A nossa abordagem está de acordo com o ideal de eliminação do 4WHP proposto neste protocolo. Além disso a nossa solução vai mais à frente, especificando como os APs obtêm o material de chaves tornando a nossa abordagem independente da implementação. Estas chaves são distribuídas sem serem necessárias alterações nos elementos da rede para suportar protocolos inter-AP. Cada AP no domínio de mobilidade obtém novo material de chaves, nomeadamente a PMK, que é única para cada par MN-AP e não é partilhada por nenhum outro AP tal como no 802.11r. Desta forma, APs sob influência maliciosa não influenciam a segurança de outros APs do domínio de mobilidade. Por fim nós conferimos autenticação dos protocolos de reassociação enquanto o 802.11r não o faz.

Capítulo 4

Reautenticação 802.1X Durante a Fase de Perscrutação

O objectivo geral deste trabalho consistiu em melhorar o mais possível a experiência de mobilidade de um dispositivo móvel mantendo um nível de segurança equivalente ao protocolo 802.11i. A solução aqui apresentada constituí uma solução para a transição rápida de MNs entre APs num cenário Intra-Domínio.

Num trabalho anterior [1], foi aplicada uma solução de reautenticação 802.1X que utilizava os protocolos de Autenticação e Associação mediante o uso das tramas características, nomeadamente *Authentication Request/Response* e *Association Request/Response*. Neste trabalho será apresentada uma outra aproximação, usando as tramas características do protocolo de perscrutação do meio, concretamente as tramas de *Probe Request/Response* em conjunto com as tramas de *Association Request/Response*.

A principal vantagem desta alteração reside no facto de se tornar possível o uso de sessões de perscrutação do meio feitas de forma proactiva para realizar parte das operações de reautenticação. Mais ainda podem criar-se condições para uma reautenticação rápida em múltiplos APs, o que facilita a tomada de decisão relativamente à transição para inúmeros APs.

Na primeira secção deste capítulo é descrito o trabalho em que nos baseámos [1] que utiliza tramas dos protocolos de autenticação e associação 802.11 para tratar da criação de SAs (Security Associations), nomeadamente tramas *Authentication Request/Response*. Na segunda parte é exposta uma abstracção do protocolo apresentado em [1], onde não se faz qualquer alusão ao mapeamento do mesmo em tramas 802.11 específicas de forma a mapear os valores que devem ser trocados entre os diferentes intervenientes do processo. Por fim, é exposta a nossa abordagem e explicado o nosso novo protocolo.

4.1 - Protocolo de Reautenticação 802.1X Proposto em [1]

A reautenticação implica comunicação por partes dos clientes móveis com os pontos de acesso que disponibilizam os recursos necessários para aceder aos serviços de uma rede sem fios. Essa comunicação é efectuada entre o cliente e o AP com o qual se pretende que seja criada a Associação de Segurança.

Para conseguir uma reautenticação rápida 802.1X bem sucedida é necessário garantir os seguintes requisitos:

- Instalar uma chave PMK renovada tanto no MN como no AP;
- Usar a PMK e dois nonces para produzir a PTK;
- Confirmar mutuamente o conhecimento comum da PTK;
- Trocar capacidades RSNIE autenticadas;
- Enviar uma GTK confidencial do AP ao MN.

4.1.1 - Serviço de Reautenticação (*Reauthentication Service*, RS)

O RS é um serviço implementado para tratar as reautenticações rápidas 802.1X. Seguindo a terminologia do 802.11r, o RS é o serviço que possibilita a definição de um *Domínio de* Mobilidad (MD) e*,* constituído por todos os APs com informação que lhes permite chegar e interagir seguramente com o RS para tratar pedidos de reautenticação dos MNs. O RS substitui o AS (*Authentication Service*) para as reautenticações. Considera-se que em cada domínio gerido por um AS existe um RS que recebe material de reautenticação secreto do AS. Assume-se que o RS é capaz de autenticar mensagens do AS e estas apenas podem ser processadas pelo RS.

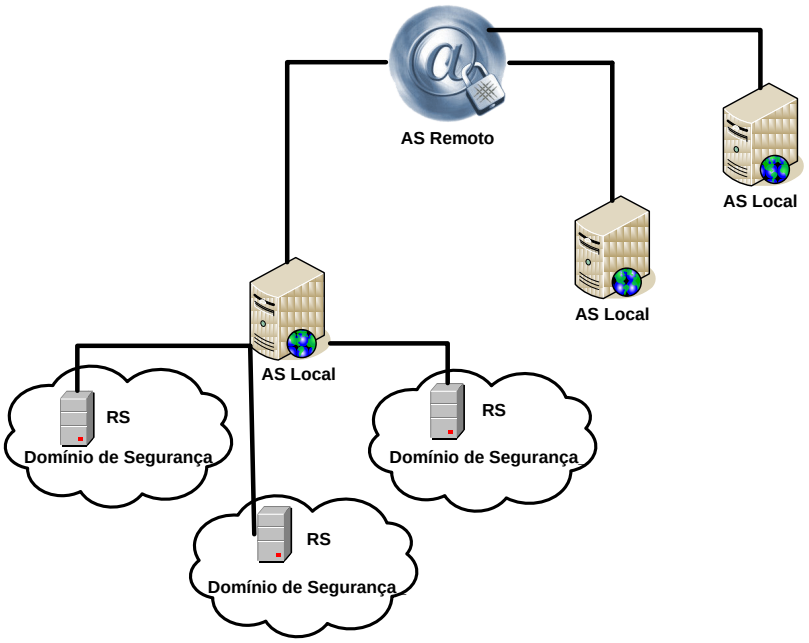

Figura 4.1: Representação de uma hierarquia de Serviços de Autenticação com a inclusão do novo Serviço de Reautenticação (RS).

O RS pode ser implementado de duas formas, ou como parte do AS local ou como um servidor HOKEY independente. Note-se que se o RS fizer parte do AS local, a comunicação entre eles não implica overhead de gestão adicional.

Os APs usam o RS em vez do AS para tratar os pedidos de reautenticação dos MNs. Assumem-se existir Associações de Segurança entre todos os APs e o RS similares àquelas entre os APs e o AS local. Estas *SAs* são essenciais para garantir a confidencialidade das chaves provenientes do AS local para o RS e deste para os APs para efectuar a autenticação de mensagens. Usando a terminologia do 802.11r, o Domínio de Mobilidade de um MN é o conjunto de APs que têm uma *SA* com o RS local.

4.1.2 - Autenticação 802.1X inicial

Para o novo protocolo de reautenticação assume-se que algum material de segurança foi produzido a priori através de uma autenticação 802.1X completa envolvendo o MN, o primeiro AP do domínio de segurança a que esse MN se associou, o AS que possibilitou a autenticação mútua e o RS que terá o seu papel em reautenticações futuras. Esse material consiste numa chave **RK** (*Reauthentication Key*) e um identificador único, **SDP** (*STA Digital Pseudonym*). Estes dois novos valores são calculados pelo *suplicante* e pelo AS durante a primeira autenticação 802.1X completa. São depois transferidos para o RS.

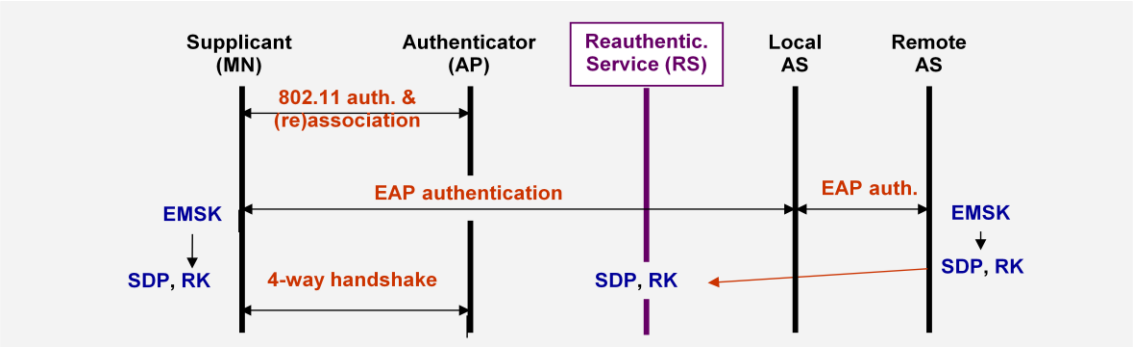

Figura 4.2: RS, integração com a arquitectura 802.1X e material secreto (SDP e RK) transferido do AS local para o RS depois de executado com êxito o protocolo EAP.

Após uma autenticação 802.1X, o suplicante partilha uma EMSK com o AS que o autenticou. Será usada esta chave na derivação da chave RK e do SDP tal como se apresenta:

RK = PRF-X (EMSK, "802.1X Reauthentication")

SDP = PRF-X (RK, ID)

onde PRF-X(Y) representa os primeiros **X** bits computados sobre **Y** por uma função pseudo-aleatória de expansão de chaves definidas em [42] e onde o ID é a identificação do suplicante fornecida pelo AS durante a sua autenticação. Segundo [3], a RK é uma chave raiz específica do domínio e o SDP é uma chave raiz de uso específico específica do domínio.

A RK será usada para gerar uma nova PMK para cada pedido de reautenticação marcado com um SDP. Usa-se o SDP em vez do endereço MAC para identificar univocamente uma sessão autenticada de um MN. A razão para tal reside no facto de o SDP não estar passível de ser falsificado por um atacante, pois deriva da EMSK, enquanto o endereço MAC é facilmente falsificado. Assim um atacante não pode lançar o ataque de endereços falsificados (spoofing) para instalar uma nova RK no RS de um MN vítima.

Dadas as *SAs* descritas, a transferência do SDP e RK do AS para o RS é tão protegida, em termos de secretismo e controlo de integridade, quanto a transferência da MSK do mesmo AS para um AP.

A reautenticação é incluída em tramas de gestão 802.11. A versão original deste protocolo usa protocolos de autenticação e associação tal como se apresenta na figura seguinte.

4.1.3 - Protocolo de Reautenticação

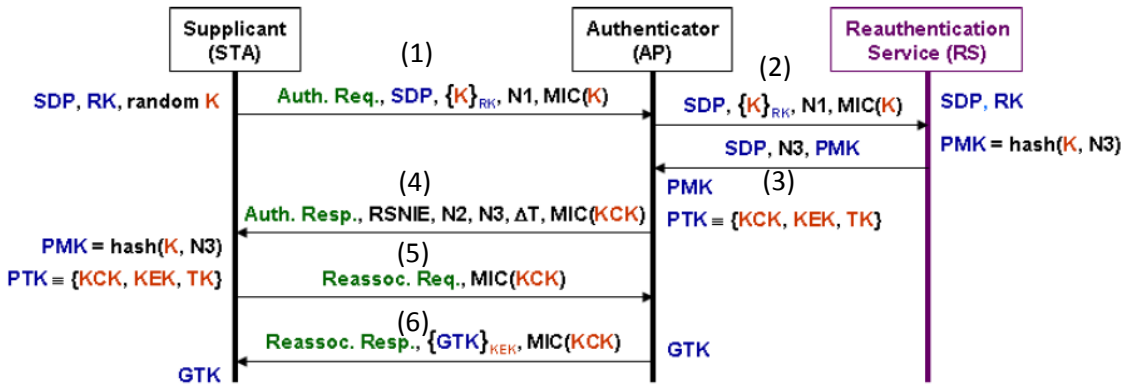

Figura 4.3: Protocolo de Reautenticação 802.1X.

MN →AP	Authentication Request, **Reauthentication Request**
MN←AP	Authentication Response, RSNIE, Reauthentication Response (sem GTK)
MN →AP	Association Request, Reauthentication Confirmation
MN←AP	Association Response, Reauthentication Response (apenas GTK)

Tabela 4.1: Mensagens de reautenticação trocadas entre os diferentes intervenientes.

Os conteúdos extra inerentes á reautenticação 802.1X podem ser adicionadas às mensagens de autenticação e (re)associação 802.11 usando novos IE.

Primeiro o MN gera o nonce N1 e uma chave aleatória K. Depois o MN envia um pedido de reautenticação ao AP contendo o seu SDP, K cifrado com a chave RK, o nonce N1 e um MIC computado com a chave K (mensagem 1). O AP encaminha todos estes valores ao RS, que usa o SDP para encontrar a chave RK correspondente ao MN (mensagem 2). Sabendo a RK, extrai a K, valida o MIC e, se válido, gera o nonce N3, calcula o hash com k e N3 para produzir a PMK e envia esta chave ao AP juntamente com o N3 identificado com o SDP correspondente ao MN que emitiu o pedido (mensagem 3).

Na sequência desta resposta do RS, o AP gera um nonce N2 e calcula a PTK com este, o N1 recebido do MN e a chave PMK recebida do RS. Note-se que o AP deve guardar o N1 aquando a recepção do pedido recebido do MN para poder a esta altura derivar a chave PTK. A resposta ao pedido de autenticação enviado do AP ao MN contém dois nonces, N2 e N3, o primeiro gerado pelo AP e o segundo gerado pelo RS e um MIC computado com a KCK (componente da PTK) identificada na Figura 4.3 pela mensagem 4. O MN usa o N3 e a sua chave K inicial para calcular a PMK e com esta, conjuntamente com N1 e N2, é calculada a PTK partilhada como no protocolo 802.1X. A KCK é então usada para autenticar a mensagem recebida por validação do MIC.

A mensagem final de confirmação da Reautenticação é enviada na mensagem de *Reauthentication Request* que prova ao AP que o MN conhece a PTK (mensagem 5). Para responder ao pedido de reassociação do MN o AP envia o *Reassociation Response* (mensagem 6) onde inclui a chave de cifra GTK cifrada com a chave KEK para poder apenas ser descodificada pelo MN a que se destina a mensagem, e também se controla a integridade da mensagem com um MIC.

No final do protocolo, o MN e o AP partilham uma nova PTK. A sua frescura é assegurada por valores aleatórios e nonces fornecidos pelos três intervenientes do protocolo: K e N1 (do MN), N2 (do AP) e N3 (do RS).

No final da reautenticação, tanto o MN como o AP estão seguros que o parceiro conhece as novas chaves (PMK, PTK e GTK) porque as mensagens de *Reauthentication Respo*nse e *Reassociation Response* são ambas autenticadas com a chave KCK. Assim o MN pode ter a certeza que o AP é genuíno, caso contrário este não poderia ter recebido a PMK do RS.

Apesar de o AP precisar receber a confirmação de reautenticação para se convencer que o MN sabe a PMK e PTK, o MN pode ser considerado autêntico assim que o RS faz essa confirmação ao AP. Este aspecto é relevante para a implementação deste protocolo e gestão da máquina de estados do 802.11.

O valor **ΔT** no *Authentication Response* indica o período que o AP irá manter o contexto de segurança (estado autenticado e chaves PMK e PTK secretas) negociado com o MN. Se o MN não se associar com o AP antes de ΔT, o AP remove o contexto de segurança sem qualquer aviso. Este mecanismo evita que o AP fique sobrelotado de *SAs* enquanto dá ao MN algum feedback sobre o tempo de armazenamento do contexto. Os APs são livres de gerir os contextos na medida em que pode atribuir diferentes tempos de caching desse contexto.

Um MN deve usar este protocolo de reautenticação com todos os APs vizinhos existentes nas proximidades.

4.2 - Abstracção do Protocolo Apresentado em [1]

Excluindo a imposição do trabalho desenvolvido em [1] de usar especificamente os protocolos de autenticação e associação para a criação das *SAs*, obtém-se uma representação abstracta apresentada de seguida, onde apenas se enuncia quais os valores que devem ser trocados, de alguma forma, entre os diferentes intervenientes do processo. Sabendo quais as trocas que devem ter lugar durante o protocolo podemos decidir quando e como enviar cada uma destas mensagens independentemente e escolher qual a melhor abordagem para cada um dos passos deste percurso protocolar.

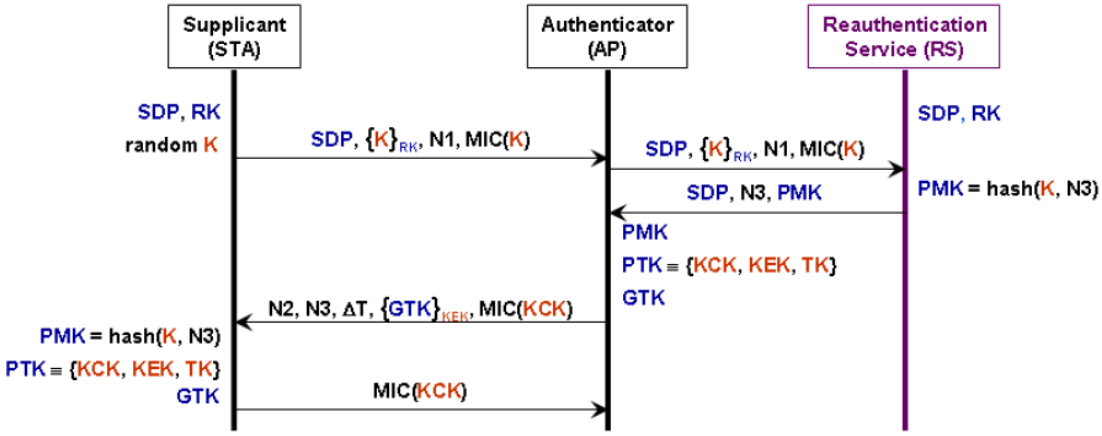

Figura 4.4: Abstracção ao protocolo de Reautenticação 802.1X definido em [1].

Mensagens do Protocolo de Reautenticação 802.1X		
Fluxo	**Tipo**	**Conteúdo**
MN→AP→RS	Pedido	SDP, $\{K\}_{RK}$, N1, MIC(K)
AP←RS MN←AP	Resposta	SDP, N3, PMK N2, N3, $\{GTK\}_{KEK}$, MIC(KCK)
MN→AP	Confirmação	MIC(KCK)

Tabela 4.2: Mensagens do Protocolo de Reautenticação 802.1X.

A Figura 4.4 ilustra o protocolo de reautenticação demonstrando os valores trocados entre os diversos intervenientes no processo sem fazer menção ao protocolo específico que leva a cabo essa troca de valores. O número de mensagens muda e o protocolo deixa de estar ligado às mensagens 802.11 usadas pelo autor de [1].

Separamo-nos da imposição de utilização dos protocolos de autenticação e associação. Referem-se apenas a existência de um pedido de reautenticação enviado do MN ao AP, o encaminhamento desse pedido ao RS e a resposta do RS até ao MN. Fica criada a SA pela troca de todos os elementos e chaves necessários à criação das chaves de comunicação. Por fim é necessário confirmar a associação do MN ao AP escolhido através da confirmação. Nessa confirmação o MN prova a sua legitimidade ao AP conseguindo a associação.

Fica assim exposta a abordagem básica do protocolo que deverá, de alguma forma, ser efectuada por meio de um qualquer protocolo de comunicação entre os intervenientes, para conseguir atingir o mesmo objectivo pretendido em [1]: *handoff* rápido e seguro.

4.3 - Reestruturação do Protocolo: Implementação do Protocolo Usando Probing e Associação

Na implementação original do protocolo de Reautenticação 802.1X proposto pelo autor Rodolphe Marques em [1], é utilizado o protocolo de autenticação como base da reautenticação 802.1X. Porém, vários benefícios podem ser conseguidos mediante o uso do protocolo de probing como base da difusão de mensagens de reautenticação pela rede. Nesta secção apresenta-se a esta nova implementação.

Comparando com o protótipo do protocolo de reautenticação proposto na abstracção apresentada na secção 4.3:

• O *Reauthentication Response* foi dividido em duas mensagens: *Probe Response* e *Association Response*. Isto acontece porque as chaves GTK devem ser distribuídas preferencialmente aquando a associação.

• O *Reauthentication Response* transporta um valor RSNIE extra. A única trama que transporta naturalmente este RSNIE é a trama de *Association Request*. Como é preciso uma resposta por parte do AP com o seu RSNIE, foi possível adicioná-lo a um *Probe Response*. Foi escolhido o *Probe Response* para antecipar a detecção de possíveis incompatibilidades entre o MN e o AP o mais cedo possível e não na altura da associação.

4.3.1 - Reautenticação 802.1X usando Probing e Associação.

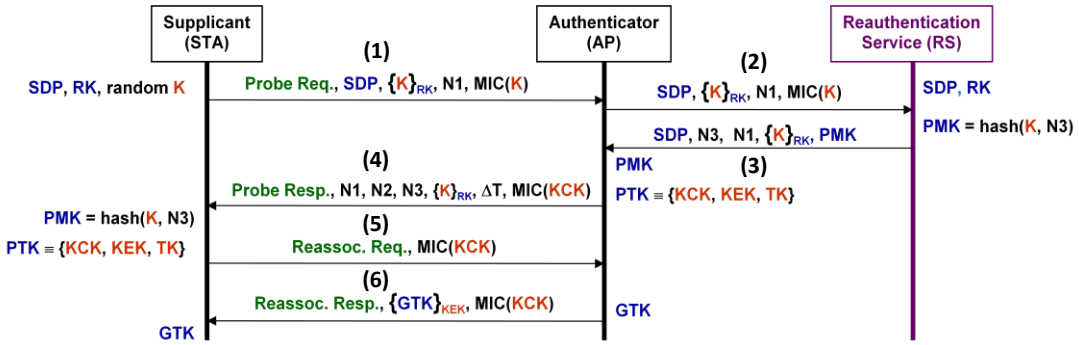

Figura 4.5: Mensagens trocadas durante a reautenticação para estabelecimento de *SAs* de um MN com APs na sua área de abrangência utilizando os protocolos de probing e associação.

Nesta secção são explicadas as modificações nos parâmetros e conteúdos das mensagens que foram efectuadas sobre o protocolo de reautenticação por meio dos protocolos de autenticação e associação. A principal diferença desta implementação baseia-se na difusão dos valores relevantes à reautenticação sobre a forma de IE nas tramas de *Probe Request/Response* do protocolo de

perscrutação activa do meio. Além disto também são introduzidas alterações nos valores constantes das mensagens. Como todo o conteúdo da mensagem de pedido de reautenticação que chega ao AP é encaminhado ao RS sem que seja guardada informação nenhuma relativa a este pedido ao nível do AP, torna-se necessário, na situação em que o RS aceita a criação da *SA,* que o RS reenvie ao AP o nonce N1 e a chave K cifrada com RK. N1 servirá para o cálculo da PTK e K cifrada para incluir na mensagem de resposta de reautenticação ao MN (mensagem 4). Nesta mensagem de resposta também é introduzido o N1. A inclusão do N1 e K cifrada na mensagem 4 deve-se ao facto de o MN renovar estes valores a cada pedido e não os registar, sendo necessário enunciá-los para criação da *SA* ao nível do MN com os valores correctos correspondentes aos existentes no AP respectivo. Estes processos são explicados de seguida.

4.3.2 - Criação das Associações de Segurança

Tendo em conta a Figura 4.5, descreve-se agora a criação das *SAs* entre o MN e cada AP ao seu alcance.

4.3.2.1 - Mensagem 1

Independentemente das políticas de gestão de mobilidade que sejam estabelecidas, o MN terá de criar as *SAs* com os APs na vizinhança num qualquer instante. Inicia-se então o envio de mensagens *Reauthentication Request* em tramas *Probe Request* aos APs de cada canal, onde consta o SDP criado na primeira autenticação 802.1X do MN, a chave K criada por uma função de geração de bytes aleatórios cifrada com a RK, o contador N1 e o MIC da mensagem completa codificado com a chave K com uma função de cifra (e.g. HMAC_SH1).

O valor N1 deixa de ser um nonce e passa agora a ser um contador. Este valor é actualizado pelo MN a cada trama de *Probe Request* de cada sessão de perscrutação que percorre todos os canais de comunicação da rede sem fios. Desta forma consegue-se que seja instalado um N1 diferente em todos os APs de acesso à rede na vizinhança do MN e impede que mensagens sejam repetidas por APs controlados por um atacante pois o RS mantém um registo do último valor de N1 recebido, não podendo ser recebido novamente esse N1 ou um N1 inferior. Para garantir uma chave PMK diferente em todos os APs varia-se a chave K por cada trama de *Probe Request* enviada numa sessão de perscrutação. N2 e N3 continuam a ser nonces.

4.3.3.2 - Mensagem 2

Um AP, com a recepção da mensagem 1, procura na sua cache indexada por SDP pela existência de uma resposta destinada ao MN. Será criada uma resposta dependendo da situação actual do processamento do pedido, que pode convergir para duas situações:

• Caso já exista essa resposta (o que significa que já ocorreu a comunicação com o RS despoletada por um pedido anterior), o AP responde de imediato com os valores necessários na mensagem 4 como apresentado na figura 4.5.

• Na eventualidade de não existir ainda uma resposta (o que significa que nunca foi realizado este pedido no passado ou, apesar de já ter existido um pedido no passado, ainda não foi recebida uma resposta do RS para este MN em particular) o AP responde com um *Probe Response* comum como descrito no standard 802.11, e encaminha o pedido ao RS (mensagem 2).

Na situação de recepção do pedido de reautenticação pela primeira vez, em que não existe qualquer registo de uma resposta pronta para o MN, o AP encaminha a mensagem recebida do MN para o RS encapsulada na forma de VP (*Value Pair*) numa mensagem RADIUS, do tipo *Authentication Request* (mensagem 2). Adicionalmente envia ao MN um Probe Response standard com a diferença de conter o IE que informa que o AP possibilita o protocolo de reautenticação. Com esta informação o MN regista que deste AP será possível extrair uma resposta no futuro. Com esta capacidade adicional, o MN poderá pedir a resposta individualmente sem ter de efectuar a perscrutação activa completa. Nenhum processamento adicional irá ter lugar ao nível do AP a esta altura do processo.

4.3.2.3 - Mensagem 3

Ao receber a mensagem *Authentication Request* (mensagem 2), o RS procura pelo VP de reautenticação no conteúdo da mensagem recebida. Ao encontrar os valores referentes ao pedido do MN que pede a reautenticação, o RS inicia o processo de verificação de integridade e confidencialidade:

• Em primeiro lugar o RS procura, na sua base de dados, a chave RK respectiva ao SDP recebido na mensagem proveniente do AP. De seguida é decifrada a chave K recebida utilizando a chave RK.

• A autenticidade é confirmada pela validação do MIC(K). Se for validado este MIC fica confirmada a integridade da mensagem recebida comprovando ao RS que o MN é legítimo e que não ouve alteração ou repetição da mensagem recebida.

• É também levada a cabo a verificação do valor de N1 e registado o último valor recebido. Para que o pedido seja considerado legítimo o valor de N1 recebido não pode ser igual ou inferior ao último valor de N1 registado.

• É então altura de criar a resposta para responder ao AP. Para tal é necessário criar o nonce N3 por meio de uma função pseudo-aleatória. Com este e com a chave K é criada a chave PMK implementando uma função de hash (e.g. SHA256). A resposta é construída onde constam o SDP do MN que originou o pedido, o N3 específico do pedido, a PMK para a possível sessão autenticada que possa vir a ser estabelecida com o AP que encaminhou o

pedido, a chave K usada no cálculo da PMK cifrada com RK e o N1 do pedido do MN para que o AP possa calcular a PTK. A mensagem é encapsulada na forma de VP de uma mensagem de *Authentication Response* e enviada ao AP que requisitou este processamento.

4.3.2.4 - Mensagem 4

O AP recebe a mensagem 3 do RS e procura pelo VP do protocolo de reautenticação. Como se assume a existência de *SAs* entre o AP e o RS, a mensagem não contém qualquer reforço de segurança por parte do novo protocolo. Após obter os valores da mensagem proveniente do RS, o AP cria a entrada na cache para a SA com o MN. Da mensagem recebida o AP extrai a PMK, e juntamente com o N1 também constante nessa mensagem e com o N2 obtido de uma função pseudo-aleatória e uma string "802.1X authentication", calcula a chave PTK por uma função de cifra PRF-X. Esta é a chave que resultaria do 4WHP num processo de autenticação 802.1X comum. Como dito anteriormente, esta PTK é decomposta em três outras chaves, a KCK, KEK e TK. A PMK e a PTK são armazenadas pelo AP, juntamente com o N1 e SDP correspondentes a este MN específico. O AP calcula o tempo que tenciona manter a SA activa e acciona o timer ΔT.

Para pedir a confirmação da criação da SA com um AP, o MN repete o envio de pedidos da primeira mensagem do protocolo de reautenticação. Ao receber este pedido, o AP verifica a existência de uma entrada na sua cache para o SDP respectivo ao MN. Se esta entrada existe em cache, o AP constrói a resposta (mensagem 4) onde irão constar os valores de N1, N2, N3 e chave K cifrada com RK com que foram calculadas as chaves PMK e PTK para criação da SA do AP com o MN. Nesta mensagem de resposta também consta o ΔT, atribuído pelo AP, que sinaliza ao MN o tempo que esse AP irá manter o contexto em cache.

Para associar esta resposta ao pedido do MN actualmente enviado, o MIC criado desta mensagem também incluí o N1 enviado ao AP neste segundo pedido mas que não foi usado para cálculo nenhum.

O MN recebe então a resposta de um determinado AP num determinado canal, com o IE do protocolo de Reautenticação 802.1X. Na verificação do conteúdo, o MN obtém o N1 e chave K usado na criação da SA, N2 do AP e o N3 do RS. Com K e N3, o MN calcula a PMK da mesma forma que esta é processada no RS. De seguida, com esta PMK, N1 gerado no início do processo de reautenticação e N2 recebido do AP, o MN calcula a PTK equivalente à do AP. Falta portanto verificar a integridade da mensagem por validação do MIC, que é calculado sobre a mensagem recebida e usando a KCK (componente da PTK calculada no MN). Após validação bem sucedida o MN guarda em cache o MAC do AP juntamente com as chaves PMK e PTK que serão usadas se o MN transferir a sessão para este AP específico por meio do processo de *handoff*. É também registado e controlado o timer ΔT.

4.3.3 - Processo de (Re)Associação

De seguida é explicado como é implementado o protocolo de (re)associação, aplicado na efectivação do *handoff*, afiliando um MN que transitou de AP com o AP que este decidiu ser o mais apropriado para a transição, contando que a reautenticação por meio do protocolo de probing já teve lugar anteriormente, e que a SA ainda está disponível no AP.

O nosso novo protocolo de associação 802.11 tem duas diferenças em relação à versão original tal como descrita pelo standard IEEE 802.11: autentica valores RSNIE trocados e envia a GTK do AP ao MN. Isto é conseguido por adição de alguns elementos nas tramas de *Reassociation Request/Response* como demonstrado na figura 4.5.

MN →AP	**Reassociation Request**, MIC(KCK)
MN←AP	**Reassociation Response**, $\{GTK\}_{KEK}$, MIC(KCK)

Tabela 4.3: Mensagens de (re)associação trocadas entre o MN e o AP.

Os MIC nas tramas de *Request* e *Response* autenticam os valores RSNIE trocados nas tramas. Este MIC prova ao AP que o MN conhece efectivamente as chaves PMK e PTK.

O MN, quando determina que é altura ideal para efectuar o *handoff*, envia uma mensagem de *Reassociation Request* (mensagem 5) onde consta um IE contendo um MIC processado sobre toda a mensagem e computado com a KCK referente à PTK calculada pelo MN. Desta forma o MN consegue provar ao AP que conseguiu derivar a chave PTK por demonstrar conhecer a KCK.

O AP valida o MIC recebido e fica convencido que o MN é genuíno e conhecido pelo RS. Todo o processamento necessário para associação do MN é levado a cabo pelo AP.

No final apenas falta confirmar ao MN o sucesso ou fracasso do pedido de associação. Para tal o AP envia o *Reassociation Response* (mensagem 6) ao MN que pediu a reassociação. Nesta trama consta a chave GTK cifrada com a KEK e um MIC criado com KCK para autenticação da mensagem ao nível do MN.

No MN, depois de validada a mensagem recebida do AP, são instaladas as chaves necessárias à comunicação com este, ou seja, a PTK e a GTK. Partindo deste momento o MN encontra-se associado com o novo AP.

Depois de associado ao novo AP o MN guarda as SAs anteriores para eventuais transições seguintes. Para tal continua a actualizar o timer ΔT com cada AP, tal como explicado na secção seguinte.

4.3.4 - Reauthentication Refresh

Para evitar a sobrecarga do AP e desembaraçar a sua cache de pedidos passados e desnecessários, é implementado um ΔT. Como descrito anteriormente, este valor simboliza o tempo durante o qual o AP irá manter o estado de reautenticação do MN, e é notificado ao cliente no *Reauthentication Response* sobre a forma IE incluído num *Probe Response*. A contagem decrescente do ΔT é iniciada quando o AP recebe a resposta positiva proveniente do RS.

Se o ΔT expira, o estado de reautenticação do MN no AP é perdido e será necessário efectuar o pedido ao RS novamente na eventualidade de o MN repetir o pedido de reautenticação com esse AP. O MN, com esta informação, poderá gerir as *SAs* em vigor a cada momento no domínio de mobilidade em que se encontra.

Após uma reassociação bem sucedida, o MN não notifica os outros APs na vizinhança sobre o seu estado de associação, pelo que as *SAs* com eles criadas não são removidas antes que o tempo anunciado em ΔT expire.

Desta necessidade de manter *SAs* presentes e passadas activas durante mais tempo para beneficiar o processo de *handoff*, é desenvolvido um novo tipo de mensagem, o *Reauthentication Refresh*. Mais uma vez é identificada pelo primeiro byte da secção de dados que diferencia os diferentes IEs deste protocolo. O MN transmite esta mensagem em broadcast no canal onde o AP a que se destina o pedido se encontra. Esta mensagem permite reiniciar o contador de ΔT para prolongar a duração da *SA*.

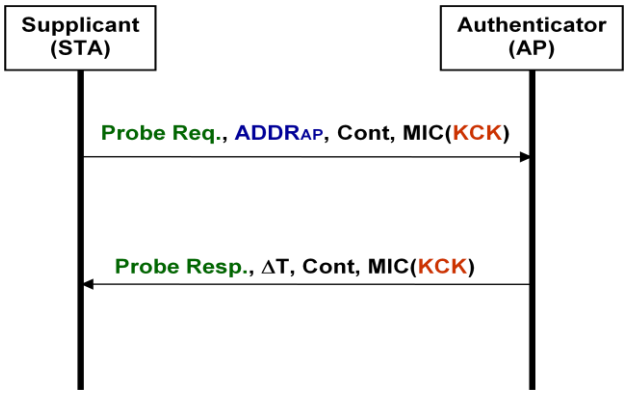

Figura 4.6: Sequência de Mensagens Reauthentication Request.

O *Reauthentication Refresh* é usado em duas situações distintas:

• Na fase de reautenticação, quando o MN se encontra a efectuar a criação de *SAs* com os APs na vizinhança. Contudo tais associações podem não vir a ser utilizadas durante o tempo limite distinto para cada uma evidenciado pelo ΔT com que se encontram marcadas. Se é intenção do MN manter estas associações activas até à efectivação do *handoff*, será transmitida uma mensagem de

Reauthentication Refresh para actualizar o valor de ΔT. Com isto é poupado um novo processo de autenticação com o RS para voltar a estabelecer a *SA*.

• No instante em que um MN muda a sua associação e passa a ser servido por outro AP marca o início de uma nova sessão de *handoff*, onde novas chaves e valores são calculados para tratar o processo de reautenticação. Contudo, para poupar recursos e minimizar a sobrecarga no RS, tornou-se possível manter *SAs* anteriores. Assim, em APs onde as *SAs* já se encontram estabelecidas, apenas se torna necessário actualizar o tempo que estas associações se mantêm activas. Tal é feito com as tramas de *Reauthentication Refresh*. Deixa de ser necessário recalcular novas chaves PMK e PTK para essa *SA* uma vez que já foram calculadas algures no passado e ainda se encontram activas tanto no MN como nos APs.

Para além de reduzir a carga no RS, o facto de estas mensagens serem enviadas apenas no canal onde se encontra a operar o AP alvo reduz a carga de rede característica de uma perscrutação completa bem como diminui substancialmente o tempo que o MN despende nesta fase, isto porque não necessita de difundir a mensagem em todos os canais mas sim em apenas um para cada mensagem.

4.3.4.1 - Funcionamento do Protocolo de Refresh

Para actualizar o ΔT, o MN envia um *Probe Request* com o IE de Refresh contendo o endereço do AP destino, um contador "Cont" e um MIC calculado sobre toda a mensagem cifrado com KCK, componente da PTK criada no MN aquando a reautenticação que estabeleceu a SA. O endereço MAC do AP consta da mensagem para identificar o AP específico ao qual se destina o pedido. Como o *Probe Request* é enviado em broadcast num canal específico, APs que recebem esta mensagem tratam do seu descarte quando, aquando o processamento da mensagem, identificam que esta não se dirige a si. Pelo endereço MAC de origem presente no cabeçalho do *Probe Request*, o AP identifica o MN que enviou o pedido e efectua a procura da *SA* respectiva presente na sua cache.

O contador existente na mensagem de *Reauthentication Refresh* é incrementado a cada par de mensagens de pedido e resposta de refresh de uma SA, ou seja, quando um MN envia um Reauthentication Refresh inclui um contador, o AP processa o pedido, armazena na sua cache o valor do contador apenso à informação relativa ao MN em questão, e responde com o Probe response respectivo. Se um novo pedido de Reauthentication Request for enviado pelo MN, o contador é incrementado. Quando o AP recebe esta mensagem verifica se o contador recebido é superior ao anterior e valida o pedido actualizando o ΔT respondendo ao MN. Este mecanismo de utilização de contadores traduz-se numa medida de segurança: se um atacante repetir a mensagem de refresh, esta é descartada pois o valor do contador é igual ou inferior ao anterior recebido. Se este valor for incrementado pelo atacante irá dar origem a uma falha de segurança identificada na altura de validação do MIC no AP. Este processo é necessário para evitar que a SA seja mantida indefinidamente, sobrecarregando a cache do AP. O MIC serve obviamente para validação da integridade do conteúdo da mensagem e também para provar ao AP que foi um MN autentico que

enviou o pedido por demonstrar que tem conhecimento da PTK criada na sessão correcta de reautenticação.

Após estas verificações o AP actualiza o ΔT permitindo que a *SA* se preserve. Para notificar o MN desta actualização, o AP responde com o *Probe Response* onde integra um novo IE no qual consta o valor de ΔT actualizado, o contador incrementado e novamente um MIC da mensagem calculado com KCK para manter a integridade da mensagem e assegurar o MN que o pedido foi processado por um AP genuíno com conhecimento das chaves PMK e PTK correctas. Se um AP não se encontra ao alcance do MN não poderá ouvir o Reauthentication Request. Nesta situação não será gerada uma resposta para o MN que descarta esta *SA* e o ΔT no AP expira sendo também removida a SA.

No nosso novo protocolo de reautenticação apenas uma trama de *Reauthentication Request* enviado por um MN genuíno poderá actualizar o valor de ΔT num AP específico.

4.3.5 - Information Element (IE) do Protocolo de Reautenticação

De modo a facultar uma forma prática e fácil de adicionar informação a tramas com uma estrutura já definida propõe-se a utilização de Information Elements. Um IE constitui uma parcela de uma trama de gestão em redes sem fios 802.11. Estes IEs desempenham a função de transferência de informação entre dispositivos, informação essa relativa a si próprios ou a protocolos em decurso. Geralmente existem vários IEs em cada trama e cada um é construído sobre a forma Tipo-Tamanho-Valor (*Type-Length-Value*, TLVs). A estrutura comum de um IE é esquematizada na Figura 4.7.

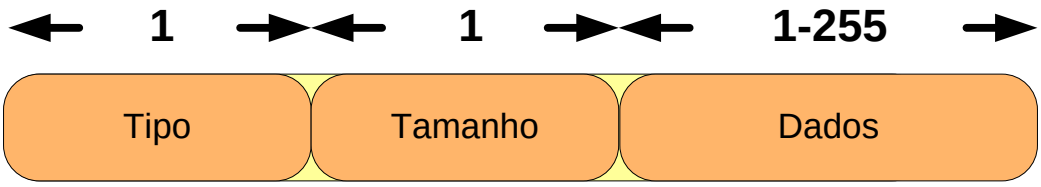

Figura 4.7: Estrutura genérica de um Information Element, tamanho dos campos apresentado em número de octetos.

Para a nossa implementação é criado um novo "tipo" de IE cujo identificador não se encontra em utilização actual. Este tipo pode ser mantido inalterado para todos os IEs do novo protocolo, quer sejam empregues em tramas de Probing, Autenticação ou Associação. O campo chamado "tamanho" regista o tamanho da secção de dados em número de octetos. No campo de "dados" estarão contidos os valores que desejamos transmitir entre os dispositivos envolvidos. Na nossa implementação o primeiro byte da secção de dados identifica qual a trama do processo de reautenticação que está a ser transmitida. Assim o mesmo IE pode ser modificado dependendo da fase do processo de reautenticação a ser executada no momento, independentemente da direcção de envio da trama ou do tipo de trama (Probe Request/Response, Association Request/Response). A presença deste tipo de

IE num *Beacon* ou *Probe Response* transmitida por um AP irá informar os dispositivos móveis que os recebem sobre a capacidade do AP para implementar este novo protocolo de reautenticação. Para o nosso protocolo são necessários os seguintes IEs:

• IE na trama de *Beacon* e *Probe Response* para sinalizar a capacidade de reautenticação do AP.

• IE no *Probe Request* do MN com os valores necessários ao pedido de criação da SA com o AP.

• IE no *Probe Response* do AP para responder com os valores oportunos para a criação da SA com o MN.

• IE nas tramas de *Association Request/Response* para confirmar a associação de um MN com um AP e efectuar a autenticação mútua entre eles com o nosso protocolo.

4.4 - Armazenamento de Informação Relevante nos MNs, APs e RS

Para conseguir que o protocolo siga o seu percurso de forma estável e eficaz devem ser implementados mecanismos de armazenamento e gestão de informação relevante. O armazenamento é efectuado pelos dispositivos, MN, AP e pelo serviço RS. A gestão é independente e levada a cabo por cada um dos dispositivos.

4.4.1 - Cache do MN

O MN organiza a sua cache de reautenticação da seguinte forma:

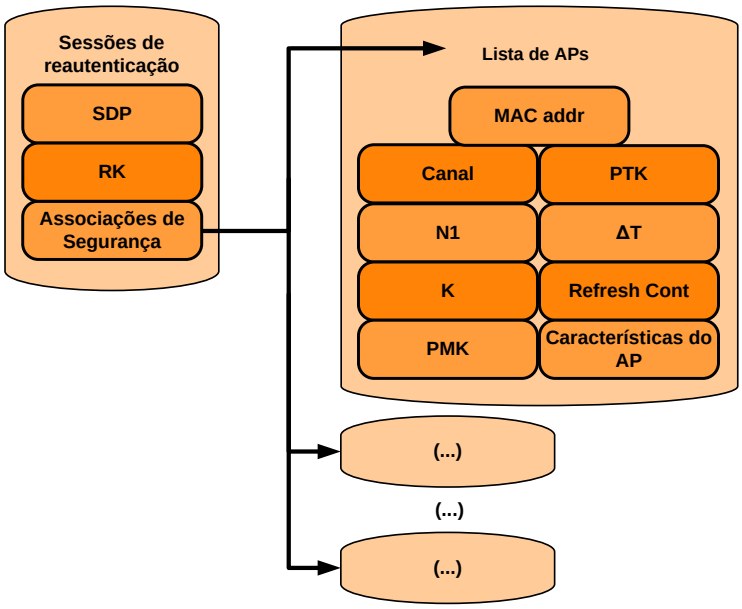

Figura 4.8: Esquema representativo da cache do MN.

Os valores principais mais relevantes do nosso protocolo são claramente o SDP e a chave RK gerados aquando a autenticação 802.1X inicial que proporciona a entrada do MN num domínio de segurança gerido por um RS.

Cada sessão de Reautenticação recruta APs que, partindo do momento que respondem ao MN com uma mensagem de r*eauthentication response*, se tornam possíveis alvos a ser considerados pelas políticas de mobilidade. Como descrito, a resposta de um AP sinaliza a criação de uma *SA* entre eles. Os parâmetros dessa SA são colocados em cache e obriga ao armazenamento dos seguintes valores:

• N1 e chave K para cálculo das chaves específicas da reassociação.

• Endereço MAC do AP e canal em que este opera. Estes dados são importantes uma vez que são necessários para difusão das mensagens de *reauthentication request* uma vez que estas são diferenciadas por canal.

• Chave PMK calculada para o AP em questão para que seja instalada no processo de (re)associação. É necessária para efectuar eventuais 4WHPs com o AP durante o funcionamento normal do protocolo 802.11 após associação com este AP.

• Chave PTK para promoção da integridade das mensagens enviadas ao AP e validação das mensagens recebidas deste (componente KCK). Também é necessária para instalação imediata aquando a (re)associação possibilitando a comunicação contínua característica do nosso protocolo.

• Timer decrescente ΔT que informa o MN sobre o tempo de vida actual da SA criada com o AP possibilitando ao MN tomar decisões sobre manutenção da AS e consequente actualização bem como decisões sobre quando deverá ser efectuado o *handoff* no intervalo de vida da AS.

• Contador de Refresh que deve constar da mensagem de Reauthentication Refresh para segurança contra ataques de repetição.

• Por fim o MN também deve guardar informação relativa às características do AP relevantes para as decisões de mobilidade. Estas características serão levadas em conta no processamento de informação que irá devolver o AP para o qual o *handoff* irá ser efectuado.

Note-se que a actualização do ΔT origina uma resposta por parte do AP que, para além de permitir actualizar este parâmetro, permite ainda a actualização da informação sobre as características do AP que podem variar consoante o MN se desloca fisicamente pelo meio.

4.4.2 - Cache do AP

O AP organiza a sua cache de reautenticação da seguinte forma:

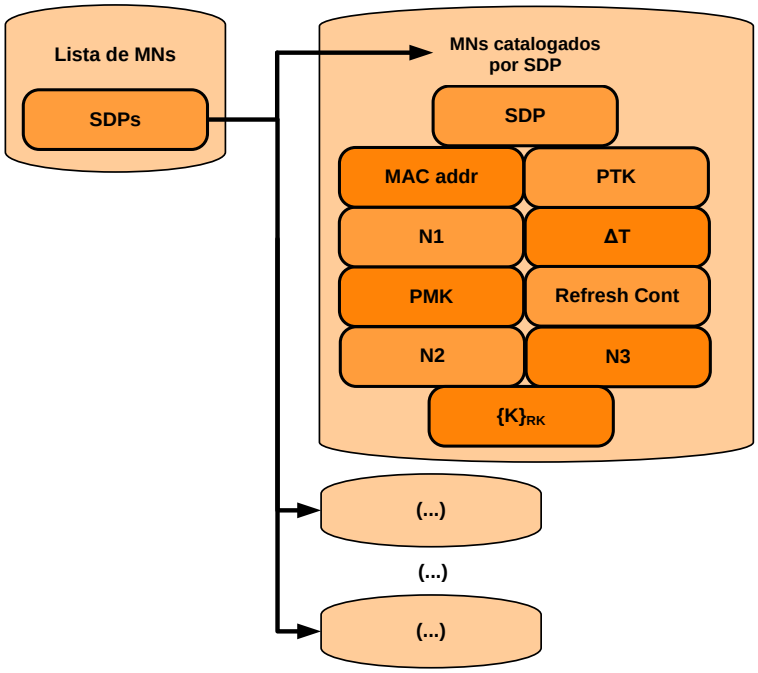

Figura 4.9: Esquema representativo da cache do AP.

O AP dispõe a sua cache na forma de lista de MNs que efectuaram um pedido de reautenticação. É adicionada a SA à cache quando o AP recebe a resposta positiva do RS com os valores necessários à reautenticação. Esta lista encontra-se catalogada por SDP pois é através deste que um MN é identificado no nosso protocolo. Cada entrada na cache do AP está dividida nos seguintes campos:

• Endereço MAC do MN para verificar a origem de mensagens de Reauthentication Refresh.

• Nonce N1 com o qual, juntamente com o N2 por ele criado e com a chave PMK recebida do RS, irá ser calculada a chave PTK para comunicação com este MN.

• Chave PMK recebida do RS e chave PTK que devem ser guardadas pelo mesmo motivo que no MN.

• Timer decrescente ΔT calculado pelo AP que lhe informa o tempo em que esta *SA* será mantida activa.

• Contador de Refresh para evitar repetição de mensagens de Reauthentication Refresh por atacantes.

• N2, N3 e K cifrada para criar a resposta (mensagem 4) para o MN quando estes valores já encontram em cache.

4.4.3 - Cache do RS

Por fim, o RS também armazena valores relevantes ao funcionamento do protocolo. A sua cache apresenta-se organizada na seguinte estrutura:

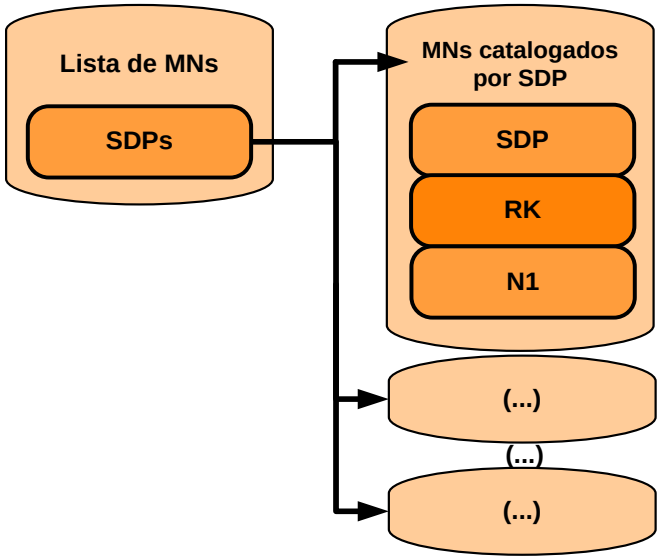

Figura 4.10: Esquema Representativo da cache do RS.

O RS simplesmente precisa armazenar os valores do SDP e RK de cada MN para autenticação dos pedidos e processamento dos valores de resposta. Adicionalmente, contando que o nonce N1 é um contador, este valor também é guardado apenso à restante informação para que possa ser realizada a verificação de segurança descrita em 4.5.5.3.

4.5 - Avaliação de Segurança

Após apresentado o protocolo convém referir quais as alterações e melhorias implementadas em termos de segurança quando comparado com o protocolo original de onde este evoluiu.

É importante garantir que MNs falaciosos não consigam obter acesso a uma rede sem fios durante o processo de *handoff*. É também essencial que um MN legítimo não se deixe enganar por um AP falso, incorrendo no risco de emitir informação sensível e confidencial a um atacante. É preciso garantir que a comunicação através do ar, que pode ser interceptada por qualquer NIC, não revele ou deixe margem para revelar a informação contida nas tramas transaccionadas.

A frescura da chave PMK é assegurada pelo MN e pelo RS, com a chave K e nonce N3, respectivamente. No novo protocolo foi instituído que a chave K é renovada a cada trama de *Probe Request*, daí a necessidade de esta ser comunicada no sentido inverso na altura da confirmação da criação da *SA*. Este reforço de segurança garante que a descoberta da chave K não influencia associações seguintes do MN.

A frescura da PTK é assegurada pela frescura da PMK, o contador N1 que será sempre diferente e o nonce N2, providenciados pelo MN e pelo AP respectivamente.

É necessário garantir que atacantes a escutar o meio não possam decifrar as chaves K e RK usada por um STA. A decifra da chave K permite a um atacante escutar a conversação, falsificar e adulterar uma sessão entre um MN e um AP que deveria ser segura. A decifra da chave RK torna-se ainda mais perigosa pois possibilita que um atacante possa personificar um MN, pode enganar um MN personificando um AP ou ainda decifrar a chave K. Se K é gerada aleatoriamente pelo MN, o atacante terá de descobrir primeiro a RK para decifrar $\{K\}_{RK}$. Contudo, os únicos valores que os atacantes vêm cifrados com a RK são os valores aleatórios de K, pelo que apenas com pesquisa exaustiva é que se torna vantajoso saber a chave RK ou K. Assim, escolher uma RK e Ks aleatórias e extensas o suficiente, com pelo menos 128 bits, deve ser suficiente para prevenir ataques de descoberta de chaves por força bruta.

Uma melhoria directamente introduzida pelo nosso protocolo de reassociação 802.1X é que tramas de *Association Request* e *Probe Response* podem ser autenticadas. Esta autenticação evita que atacantes consigam lançar ataques DoS (Denial of Service) através da emissão destas mensagens adulteradas pelo atacante. Em geral, todas as tramas trocadas entre o MN e o AP depois de um *Reauthentication Request* podem ser autenticada com um MIC calculado com a KCK. Isto também era possível com as especificações 802.11 correntes, contudo não era explorado, mas foi estendendida esta possibilidade para interacções com APs antes das associações, o que era impraticável com os modelos actuais do 802.11 e 802.1X.

Tal como explicado, o valor N1 deixa de ser um nonce e passa agora a ser um contador actualizado pelo MN a cada trama de *Probe Request*. Isto permite que seja instalado um N1 diferente em todos os APs de acesso à rede na vizinhança do MN e impede que mensagens sejam repetidas por APs controlados por um atacante. O RS mantém um registo do último valor de N1 recebido, não podendo ser recebido novamente esse N1 ou um N1 inferior. Esta medida de segurança garante que APs verdadeiros mas controlados por atacantes não possam repetir mensagens legítimas de outros APs ou mensagens encaminhadas via AP provenientes de MNs falsos e maliciosos consigam iniciar uma sessão na rede atacada. A repetição exaustiva desta mensagem ao RS iria sempre originar uma resposta. Como o RS responde com um N3 aleatório, eventualmente seria usado o mesmo nonce N3 por parte do RS para geração da chave PMK que foi originalmente usado para gerar a PMK para o par MN/AP legítimos, possibilitando ao atacante decifrar mensagens trocadas no passado entre o AP e o MN legítimos. As probabilidades de tal ocorrer tornam-se elevadas se o atacante conseguisse repetir

o mesmo pedido vezes suficientes para que seja usado o mesmo N3 originando a PMK correcta. Com o reforço de segurança conferido pelo facto de N1 ser um contador, as mensagens entre o AP e o RS não podem ser repetidas sobre pena de serem de imediato descartadas. A desvantagem desta técnica reflecte-se na situação em que existam dois APs pertencentes à mesma rede a operar no mesmo canal e ao alcance do MN. Ambos os APs recebem a mesma mensagem de reautenticação com o valor de N1 igual. No decorrer normal deste método, apenas o primeiro AP a comunicar com o RS consegue uma resposta. Tal não é grave, pois o segundo AP conseguirá tratar do pedido para o MN na sessão de perscrutação seguinte. Nas redes estruturadas actuais bem implementadas é raro encontrar dois APs com proximidade física tal que possibilite a intersecção da mesma área de abrangência a funcionar no mesmo canal de comunicação, assim sendo o problema criado por dois APs no mesmo canal e servindo a mesma rede (SSID) ao alcance de um MN é mínimo.

Para associar a resposta do AP com informação da *SA* ao pedido de criação da SA por parte do MN (situação em que já existe no AP uma resposta para um pedido anterior do MN), o MIC criado desta mensagem também incluí o N1 enviado ao AP no pedido de resposta, valor de N1 esse que não foi usado para cálculo nenhum. Desta forma o MN fica assegurado que esta mensagem foi enviada pelo AP como resposta ao seu pedido, não se tratando de uma mensagem falaciosa criada por um atacante que faria o MN acreditar existir uma SA com um AP fictício, o que iria interferir nas suas decisões de mobilidade. Isto constitui uma medida de segurança pois garante a integridade da mensagem e assegura o MN que a mensagem está a ser enviada de um AP genuíno, caso contrário não poderia ter calculado o MIC com a chave k e N1.

Além destes reforços de segurança falta referir que o novo protocolo de actualização das *SAs*, o *Probe Refresh*, aparece guarnecido de um contador para validação destas mensagens tal, como é feito com N1. Esta capacidade adicional garante que as mensagens de *Probe Refresh* não podem ser repetidas impossibilitando ataques de repetição e preservação extensiva de *SAs* sem interesse.

Capítulo 5

Conclusão

Nesta dissertação foi apresentada uma nova abordagem para conseguir reautenticações 802.1X rápidas em redes IEEE 802.11. O trabalho aborda um problema actual e relevante. De facto, a capacidade de gerir a mobilidade dos MNs é essencial para determinar a qualidade/sucesso de redes sem fios futuras. Nesta dissertação apresentou-se uma solução prática que pode ser implementada na vida real.

O método apresentado permite a criação de várias *SAs* de forma proactiva na vizinhança do MN, reduzindo o tempo handoff e melhorando os critérios e políticas de mobilidade, sendo apenas necessário avaliar qual o melhor AP e decidir se devem ou não ser mantidas determinadas *SAs* provisórias ainda activas.

As reautenticações são conseguidas antes das (re)associações e as *SAs* são criadas proactivamente com APs candidatos antes de haver transferência propriamente dita entre APs. Recupera-se desta forma o paradigma do 802.11 – primeiro autenticação, depois associação – que é crucial para a implementação dos *handoffs* rápidos. Também é possível autenticar várias interacções de controlo do protocolo 802.11, nomeadamente as interacções de probing e reassociação, o que é útil contra ataques DoS. Em termos de eficiência de handoff, foi possível remover quase todo o atraso do processo de handoff, já que este pode ser preparado proactivamente antes da reassociação. Apenas a autenticação das tramas do protocolo de associação e a distribuição da GTK adiciona um atraso extra em relação ao protocolo base de associação.

Em contraste com o trabalho proposto em [1], com a aproximação proposta nesta dissertação é possível realizar reautenticações 802.1X de forma eficiente em APs na vizinhança usando tramas de perscrutação do meio, isto é, Probe Resquest/Response. Isto permite saltar a fase de autenticação 802.11 ponto-a-ponto relativa aos APs escolhidos como futuros candidatos a servir o MN, apenas é necessário perscrutar o meio na procura de APs e, mais tarde, associar-se ao mais adequado. Criou-se ainda um protocolo adicional para gerir as *SAs* vigentes, o *Probe Refresh*, para permitir a manutenção ou rescisão destas mesmas. Por fim foi clarificado o funcionamento da cache que deve existir em cada um dos intervenientes necessária para o funcionamento do protocolo.

A abordagem seguida neste trabalho é similar ao 802.11r, contudo, existem diferenças relevantes: são distribuídas as chaves PMK, o 802.11r não o faz; as tramas de *Probe Request/Response* são autenticadas, o 802.11r não explora esta funcionalidade; limita-se o tempo para manter informação na cache do AP, o 802.11r apenas o faz para variantes de protocolos que fazem reserva de recursos que usam mais mensagens.

Convém ainda referir que a solução apresentada pode ser aplicada em conjunto com qualquer outra solução de gestão do *handoff* que atribua ao MN a responsabilidade de tratar do processo. Assim, seja usada perscrutação proactiva ou não, sejam usados NGs ou outras estruturas para indicar características ou posições de APs preferenciais, ou qualquer outra técnica, a partir do momento em que são usadas tramas de *Probe Request* podem ser estabelecidas *SAs* com os APs e colocar o nosso protocolo em funcionamento.

Como trabalho futuro será necessário efectuar intensivos testes de validação e quantificação de ganhos de desempenho da proposta feita nesta dissertação. Estes testes deverão ser o mais abrangente possível e englobar cenários em meio laboratorial e em redes em operação.

Referências

[1] Rodolphe Marques, “Security and Mobility in 802.11 Structured Networks”, Master Thesis, Departamento de Electrónica, Telecomunicações e Informática, Universidade de Aveiro, 2008.

[2] M. S. A. Mishra and W. Arbaugh. An Empirical analysis of the IEEE 802.11 MAC Layer *Handoff* Process. ACM SIGCOMM Computer Communication Review, 33(2):93{102, April 2003.

[3] Bernard Aboba, “Fast *Handoff* Issues”, IEEE 802.11-04/827r0, July 2004.

[4] IEEE 802.11, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, 1999. [2] N.R. Prasad, and A.R. Prasad, editors, *WLAN Systems and Wireless IP for Next Generation Communications*, Artech House, Norwood MA, January 2002.

[5] C. Perkins, “IP Mobility Support”, RFC 2002, IETF, Oct. 1996.

[6] H.Velayos and G. Karlsson, “Techniques to Reduce IEEE 802.11b Mac Layer Handover Time,” KunglTekniska Hogskolen, Stockholm, Sweden, Tech. Rep. TRITA-IMIT-LCN R 03:02, ISSN 1651-7717, ISRN KTH/IMIT/LCN/R-03/02-SE, April 2003.

[7] IEEE 802.11i, *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Security Enhancement*, Draft 10.0, July 2003.

[8] IEEE 802.1X, *Port-Based Network Access Control*, Oct. 2001.

[9] R. Droms, "Dynamic Host Configuration Protocol," RFC 1541, Oct. 1993.

[10] IEEE 802.11f, “Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation”, P802.11f, January 2003.

[11] L. Blunk and J. Vollbrecht, “PPP Extensible Authentication Protocol (EAP)”, RFC 2284, IETF, Mar. 1998.

[12] B. Aboba. Fast *Handoff* Issues. IEEE-03-155r0-I. IEEE 802.11 Working Group, March 2003.

[13] B. Sarikaya and X. Zheng, “CAPWAP Handover Protocol”, in *IEEE Int. Conf. on Communications (ICC’06)*, June 2006, vol. 4, pp. 1933–1938.

[14] Raymond Greenlaw and Paul Goransson, *Secure Roaming in 802.11 Networks*, Elsevier, 2007, ISBN-13 978-0-7506-8211-4.

[15] S. Mangold and L. Berlemann, “IEEE 802.11k: Improving Confidence in Radio Resource Measurements,” in *IEEE 16th International Symposium on Personal Indoor Mobile Radio PIMRC*, September 2005.

[16] Y. Bejerano R. Bhatia. Mifi: A framework for fairness and QoS assurance in current IEEE 802.11 networks with multiple access points. In *IEEE Infocom*, 2004.

[17] Y. Bejerano S. Han L. Li. Fairness and load balancing in wireless lans using association control. In *MobiCom*, 2004.

[18] H. Wu, K. Tan, Y. Zhang, and Q. Zhang, “Proactive scan: fast handoff with smart triggers for 802.11 wireless LAN,” in Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07), pp. 749–757, Anchorage, Alaska, USA, May 2007.

[19] Ganesh Venkatesan et. al., “IEEE 802 Tutorial: Video over 802.11”, in 802 Tutorials from IEEE, March 2007.

[20] V. Brik, A. Mishra, S. Banerjee, Eliminating *handoff* latencies in 802.11 WLANs using multiple radios: applications, experience, and evaluation, in: Internet Measurement Conference 2005, October, 2005.

[21] H. Wang and A.R. Prasad, “Fast Authentication for Inter-Domain Handover,” ICT 2004, Brazil, August 2004, in-print.

[22]V. Mhatre and K. Papagiannaki, “Using smart triggers for improved user performance in 802.11 wireless networks,” in *Proceedings of MobiSys*, June 2006, pp. 246–259.

[23] Gurpal Singh, Ajay Pal Singh Atwal, B.S. Sohi, "Multimedia Ready *Handoff* Technique for 802.11 Networks," adcom, pp.612-619, 15th International Conference on Advanced Computing and Communications (ADCOM 2007), 2007.

[24] Ishwar Ramani and al.: SyncScan: Practical Fast *Handoff* for 802.11 Infrastructure Networks. Proceedings of the IEEE INFOCOM Conference, Miami, March 2005.

[25] Y. Liao and L Gao, “Practical Schemes for Smooth MAC Layer *Handoff* in 802.11Wireless Networks”, WoWMoM ‘06, 2006.

[26] C.-C. Tseng et al., “*Location-based Fast Handoff* for. *802.11 Networks*,” IEEE Commun. Lett., vol. 9, no. 4,. 2005, pp. 304–06.

[27] Yazam M. Allawi et al., “Advanced *Handoff* Mechanism for Delay Sensitive Applications in IEEE 802.11 Wireless LAN”, Information and Communications University.

[28] N. Mustafa, W. Mahmood, A.A. Chaudhry, C.M. Ibrahim, "Pre-scanning and dynamic caching for fast *handoff* at MAC layer in IEEE 802.11 wireless LANs," *IEEE International Conference on Mobile Adhoc and Sensor Systems Conference*, pp. 8-122, IEEE International Conference on Mobile Adhoc and Sensor Systems Conference, 2005., 2005.

[29] Sang-Hee Park, Hye-Soo Kim, Chun-Su Park, Jae-Won Kim, Sung-Jea Ko "Selective Channel Scanning for Fast Handoff in Wireless LAN Using Neighbor Graph" PWC 2004: 194-203.

[30] M. Shin, A. Mishra and W.A. Arbaugh, "Improving the Latency of 802.11 Hand-offs using Neighbor Graphs" *Mobisys 2004 June, 2004, Boston, USA.*

[31] Mishra MSA, Arbaugh W (2004) "Context caching using neighbor graphs for fast *handoff*s in a wireless network". Technical report, University of Maryland, February 2004.

[32] **Sangheon Pack**, Hakyung Jung, Taekyoung Kwon, and Yanghee Choi, "SNC: A Selective Neighbor Caching Scheme for Fast *Handoff* in IEEE 802.11 Sem fios Networks," *ACM Mobile Computing and Communications Review (MC2R),* Vol. 9, No. 4, October 2005.

[33] A. R. Prasad and H. Wang, "Roaming key based fast handover in WLANs", in *IEEE Wireless Communications and Networking Conf. (WCNC 2005)*, Mar. 2005, vol. 3, pp. 1570-576.

[34] M. Kassab, A. Belghith, J. Bonnin, and S. Sassi, "Fast Pre-Authentication Based on Proactive Key Distribution for 802.11 Infrastructure Networks", in *1st ACM Works. on Wireless Multimedia Networking and Performance Modelling (WMuNeP'05)*, Montreal, Canada, Oct. 2005.

[35] L. Zan, J. Wang, and L. Bao, "Personal AP Protocol for Mobility Management in IEEE 802.11 Systems", in *Proc. of the 2nd Ann. Int. Conf. on Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITOUS'05)*,Washington, DC, USA, 2005, pp. 418–425, IEEE Computer Society.

[36] J. Chen, Y. Tseng, and H. Lee, "A Seamless *Handoff* Mechanism for DHCP-Based IEEE 802.11 WLANs", *IEEE Communications Letters*, vol. 11, no. 8, pp. 665–667, Aug. 2007.

[37] M. Nakhjiri and Y. Ohba, ", IETF HOKEY WG Internet-Draft, Nov. 2007, draft-ietf-hokeykey-mgm-01.

[38] R.Marin, P. J. Fernandez, and A. F. Gomez, "3-Party Approach for Fast Handover in EAP-Based Wireless Networks", in *Proc. of OTM Conferences, 2nd Int. Symp. on Information Security (IS'07)*, Vilamoura, Portugal, Nov. 2007, pp. 1734–1751, Springer, LNCS 4804.

[39] T. Clancy, M. Nakhjiri, V. Narayanan, and L. Dondeti, "Handover Key Management and Reauthentication Problem Statement ", IETF HOKEY WG Internet-Draft, Nov. 2007, draft-ietfhokey-reauth-ps-07.

[40] V. Narayanan and L. Dondeti, “EAP Extensions for EAP Re-authentication Protocol (ERP)”, IETF HOKEYWG Internet-Draft, Nov. 2007, draft-ietf-hokey-erx-08.

[41] Chung-Ming Huang and Jian-Wei Li, “An IEEE 802.11 Fast Reassociation and Pairwise Transient Key establishment Based on the Dynamic Cluster Method”, in *Works. of Computer Networks and Wireless Communications, Int. Computer Symp. (ICS 2006)*, Taipei, Taiwan, 2006.

[42] A. Mishra, Min Ho Shin, Jr. N. L. Petroni, T. C. Clancy, and W. A. Arbaugh, “Proactive key distribution using neighbor graphs”, *IEEE Wireless Communications*, vol. 11, no. 1, pp. 26–36, Feb 2004.