White Paper: Coleta de endereços MAC pela Receita Federal

Vinícius da Silveira Serafim*
vinicius@serafim.eti.br

Resumo: no dia 17 de abril de 2015, uma notícia vinculada na Internet informou que 80 mil contribuintes foram intimados pela Receita Federal do Brasil por suspeita de terem fraudado declarações do IR. Iágaro Jung, subsecretário de Fiscalização da Receita Federal, apontou o trabalho do Laboratório de Lavagem de Dinheiro da Receita Federal pela potencialização de identificação de operações irregulares. Esse laboratório desenvolve ferramentas para esse fim e uma das ferramentas utilizadas envolve a coleta de endereços IP e endereços MAC dos computadores utilizados na gravação das declarações. Neste artigo foram analisadas as formas pelas quais esses dados são obtidos, bem como a utilidade dos mesmos na identificação de computadores, usuários e contribuintes. Algumas possibilidades de cruzamento de dados foram apontadas e algumas considerações que dão margem para uma discussão jurídica foram também realizadas.

1 Introdução

No dia 17 de abril de 2015, em matéria publicada pela Agência Brasil[1], cerca de 80 mil contribuintes foram identificados pela Receita Federal do Brasil como suspeitos de fraudar declarações do IR (Imposto de Renda). Ainda conforme a matéria, o subsecretário de Fiscalização da Receita Federal, Iágaro Jung, apontou esse resultado como sendo devido à atuação do Laboratório de Lavagem de Dinheiro da Receita. Essa equipe pesquisa e desenvolve ferramentas para detectar possíveis fraudes nas declarações do IR e, entre as diversas informações coletadas por essas ferramentas, duas ganharam destaque na matéria citada: endereço IP (Internet Protocol) e endereço MAC (Medium Access Control).

A intenção da Receita Federal, também segundo a matéria citada, é "identificar os computadores de escritórios de contabilidade responsáveis por irregularidades nas declarações dos clientes." Partindo deste ponto, é realizada uma breve análise neste artigo sobre a relevância da coleta do endereço IP e do endereço MAC, bem como sobre

[1] http://agenciabrasil.ebc.com.br/economia/noticia/2015-04/receita-intima-80-mil-contribuintes-suspeitos-de-infracao-em-declaracoes-do

a forma como essa coleta é realizada. Ao final, são levantados alguns aspectos que merecem uma atenciosa avaliação jurídica.

2 Endereço IP

O endereço IP identifica o "local" na Internet a partir do qual foi enviada uma declaração. Porém o IP não identifica necessariamente um computador único na Internet, uma vez que é bastante comum o compartilhamento de um único endereço entre diversos computadores de uma mesma rede. Sendo assim, seria possível identificar a pessoa (física ou jurídica) responsável pelo uso de um determinado IP em um determinado período, mas não o computador específico utilizado para a criação da declaração.

Outro ponto relevante a ser considerado é que, de acordo com a forma como o acesso à Internet é realizado, o endereço IP de uma pessoa pode mudar frequentemente ou pode ser fixo. O primeiro caso é certamente o mais frequente, tornando mais difícil a identificação da pessoa responsável pelo uso do endereço IP. Nesse caso, a Receita teria que contar com a cooperação do provedor de acesso à Internet para saber quem usava o endereço em uma determinada janela de tempo.

Assim como há os endereços IP utilizados na Internet (chamados de IPs válidos) há também endereços para serem utilizados especificamente em redes internas (IPs inválidos)[2]. Estes últimos podem de fato identificar de forma única um computador, dentro de uma rede específica, de forma análoga ao número de uma casa em uma rua específica de uma cidade. Apenas com um IP inválido, sem saber qual a rede, a Receita teria o mesmo problema de identificar uma casa sabendo apenas o número da mesma e não conhecendo o nome da rua e mesmo o nome da cidade.

2.1 Como a coleta é feita

Quando dois computadores se comunicam através de uma rede é necessário que um saiba o endereço do outro ou a comunicação simplesmente não seria possível. Dessa forma, é trivial para um *software* servidor armazenar em seus registros de acesso os endereços IP dos clientes que nele conectaram e solicitaram ou enviaram informações. De fato, diversos *softwares* servidores têm esse como um comportamento padrão (*default*).

No caso dos endereços IP válidos, a coleta é fácil de ser feita e pode acontecer até mesmo sem que a Receita Federal tenha a intenção de realizá-la, torna-se assim perfeitamente possível saber a partir de qual endereço IP válido uma determinada declaração foi enviada.

Em se tratando de endereços IP inválidos, ou de uso exclusivamente interno, a coleta não se dá de forma tão “automática” pois esses endereços não são visíveis na Internet. Nem por isso

[2] RFC1918 disponível em http://www.rfc-editor.org/info/rfc1918

a coleta seria impossibilitada, podendo ser realizada de forma semelhante à utilizada para a coleta de endereços MAC pela Receita Federal, explicada a seguir.

3 Endereço MAC

Diferentemente de um endereço IP, que pode ser válido ou inválido na Internet, um endereço MAC é único no mundo, apesar do mesmo nunca passar de uma rede para outra. Ele é único pois a sua atribuição é controlada por uma organização, a IEEE-SA (Institute of Electrical and Electronics Engineers - Standards Association)[3]. E o fato de nunca passar de uma rede para outra, faz com que seu uso fique limitado à rede local onde se encontra o computador em uso e, portanto, jamais a Receita Federal receberia essa informação de forma automática.

O endereço MAC é atribuído não ao computador em si, mas sim as suas interfaces de rede (*wireless* ou não). Cada interface possui seu próprio e único endereço MAC, atribuído no momento de sua fabricação e por essa razão é chamado também de *endereço físico*. Cabe lembrar aqui que diversos dispositivos como *smartphones* e *tablets* são de fato computadores e possuem também suas interfaces de rede, portanto possuem endereços MAC únicos.

Como os endereços são atribuídos pela IEEE-SA aos fabricantes de equipamentos, é possível saber, também através do endereço MAC, qual é a marca do equipamento (ex. Dell, Apple, Samsung, Intel,…)

Um computador pode receber um novo endereço MAC ou perder um endereço MAC a medida que interfaces de rede são adicionadas ou removidas. Esses procedimentos, embora não muito frequentes, acontecem com maior facilidade em computadores *desktop*, sendo mais raros em *notebooks* e excepcionais em *smartphones* e *tablets*.

Existem formas relativamente fáceis (para alguém com conhecimento técnico) de alterar os endereços MAC de uma interface de rede, mas isso muito raramente é necessário.

Essas características tornam o endereço MAC mais efetivo do que o endereço IP no que diz respeito a identificar um computador (ou dispositivo) específico, inclusive sua marca, embora não seja a prova de falhas.

3.1 Como a coleta é feita

Conforme já explanado, ao contrário de um IP válido, o MAC não é enviado de forma automática para a Receita Federal. Assim, para que a receita obtenha essa informação ela tem que recorrer a outros meios. Esses meios envolvem acesso direto à rede local onde se encontra o computador a ser identificado ou, melhor ainda, ao próprio computador.

---

[3] http://standards.ieee.org/develop/regauth/oui/public.html

Uma vez que, para realizar a declaração do IR, o contribuinte deve realizar o *download* e a execução de um programa desenvolvido pela Receita Federal em um computador, a mesma obtém o acesso necessário para coletar ativamente o endereço MAC e mesmo outras informações como endereços IP inválidos (ou internos) assim como outros números de série únicos existentes no computador.

A partir dessas constatações, foram realizadas análises dos próprios programas fornecidos pela Receita Federal para verificar se de fato essa informação era coletada, pois a receita não informa o usuário sobre isso, e, se coletada, como é coletada.

3.1.1 Análise dos programas

A primeira ação realizada foi abrir um arquivo gerado pelo programa de declaração do imposto de renda pessoa física (IRPF) 2015 (extensão .DEC) e verificar se nele constava o endereço MAC de alguma interface de rede do computador onde o mesmo foi gerado. Um dos endereços MAC do computador, no formato 00:00:04:c3:de:12[4], foi encontrado logo nas primeiras linhas do arquivo .DEC, porém sem os dois pontos (:) e com todas as letras maiúsculas, seguido de 8 zeros: **000004C3DE12**00000000.

Uma vez constatado o fato de que a coleta do endereço MAC é realizada, passou-se para a análise do programa de declaração do IRPF 2015. O programa é desenvolvido em linguagem JAVA e distribuído pela Receita em seu site[5]. Após a instalação é possível encontrar o arquivo principal do programa, cujo nome é “irpf.jar”, em diretório que varia de acordo com o sistema operacional utilizado:

- Windows: C:\Arquivos de Programas RFB/IRPF2015/irpf.jar
- Mac OS X: /Applications/ProgramasRFB/IRPF2015/IRPF2015.app/Contents/Java/irpf.jar
- Ubuntu Linux: $HOME/ProgramasRFB/IRPF2015/irpf.jar

O arquivo jar (**J**ava **Ar**chive) é simplesmente um arquivo compactado contendo as classes Java, arquivos de imagem, arquivos XML e outros arquivos que são necessários para o programa. Esse arquivo pode ser facilmente descompactado e seu conteúdo inspecionado.

Entre os diversos arquivos contidos no arquivo “irpf.jar”, existe um cujo nome é “mapeamentoTxt.xml”. Nas linhas 9 e 10 desse arquivo lê-se:

```
A partir desse arquivo o componente de gravação / restauração Txt sabe quais registros
o arquivo txt a ser gravado / restaurado vai ter e quais campos cada registro vai ter.
```

[4] endereço MAC fictício.

[5] http://www.receita.fazenda.gov.br

Este arquivo portanto diz respeito ao formato do arquivo .DEC gerado pelo programa. As diversas linhas seguintes, apontam nomes de campos e descrição dos mesmos. Abaixo são destacadas as mais relevantes para a análise.

```
 56 <Campo Nome="NM_NOME"     Descricao="Nome do contribuinte"        […]
 ...
 66 <Campo Nome="NOME_SO"     Descricao="Nome do Sistema Operacional" […]
 67 <Campo Nome="VERSAO_SO"   Descricao="Numero da versao do Sistema Operacional(SO)" […]
 68 <Campo Nome="VERSAO_JVM"  Descricao="Numero da versao do Java"    […]
 ...
118 <Campo Nome="ENDERECO_MAC" Descricao="Endereço físico da estação" […]
```

A linha 56 serve apenas como um exemplo de uma informação que se espera que esteja na declaração: o nome do contribuinte. Já as linhas 66, 67 e 68 apontam para o fato de que são enviados também, como parte da declaração, o nome e versão do sistema operacional bem como a versão da plataforma Java instalada. E, por fim, na linha 118 está a indicação do armazenamento do endereço MAC do computador no qual o declaração foi gravada.

Para descobrir como o endereço MAC está sendo coletado, foi realizada a descompilação[6] das classes java do programa de declaração do IRPF 2015 (contidas no arquivo "irpf.jar) e foram procuradas algumas palavras-chave que poderiam estar relacionadas à obtenção do endereço MAC, como: *network*, *interface* e *address*.

Foram encontradas duas funções relevantes, denominadas neste artigo de "fa()" e "fb()". Ambas listadas integralmente nos anexos 1 e 2, respectivamente. A função fb() é utilizada pela função fa() e tem por objetivo determinar qual é o endereço IP do computador que está efetivamente sendo utilizado para o acesso à Internet. A forma como isso é feito é curiosa. Essa função contém uma lista de endereços de sites:

```
String[]  arrayOfString=  {"www.receita.fazenda.gov.br",  "www.google.com.br",
"www.uol.com.br", "www.globo.com", "www.terra.com.br", "www.estadao.com.br" };
```

Como pode ser notado, além do site da própria Receita, os sites do Google, UOL, Globo, Terra e do Estadão fazem parte desta lista. O que a função fb() faz é tentar se conectar em cada um desses sites na ordem em que eles aparecem na listagem e, em cada tentativa, verificar qual o endereço IP do computador (IP local) utilizado para a conexão. Assim que o endereço é obtido, as tentativas de conexão cessam e a função fb() retorna o resultado para a função fa().

O endereço IP identificado pela fb() será, em geral, um endereço inválido, ou seja, utilizado apenas em redes internas. Essa informação é utilizada pela função fa() e então é descartada, não sendo adicionada à declaração.

Já a função fa() é a que efetivamente obtém o endereço MAC. Inicialmente ela chama a fb() para obter o endereço IP da interface de rede conectada à Internet. Em seguida, se a fb() retornou

[6] Através desse processo é possível obter acesso ao código fonte de um programa.

um endereço IP, a fa() tenta obter o endereço MAC da interface de rede que possui o endereço IP encontrado. Caso contrário, a fa() obtém o endereço MAC da primeira interface de rede que ela encontrar no sistema.

As duas funções são bem simples e fáceis de serem compreendidas por alguém com conhecimento em programação de computadores.

Além do programa IRPF2015, foram também analisados: IRPF2014, IRPF2013, IRPF2012, IRPF2011, IRPF2010 e IRPF2009. Foi verificado que desde 2010 os programas de declaração do IRPF apresentavam as mesmas características de captura de endereços MAC encontradas no IRPF2015.

4 Utilidade dos endereços IP e MAC para a Receita

A Receita Federal pode facilmente obter o endereço IP válido de quem enviou uma determinada declaração. Porém, conforme foi descrito na seção 2, a utilidade dessa informação é relativa pois, por exemplo, um determinado endereço IP pode em um momento estar sendo utilizado por uma pessoa e no momento seguinte por uma outra diferente. Seria então necessária a cooperação dos provedores de acesso à Internet para que a Receita pudesse saber quais endereços foram utilizados por quais pessoas e em quais períodos.

Já o endereço MAC é fixo e, a não ser que o usuário tenha algum conhecimento técnico mais profundo, raramente será mudado. Esse endereço não permite que a Receita saiba onde se encontra o computador na Internet, mas identifica de forma única este computador. Qual é então a utilidade do endereço MAC?

Uma vez que a Receita identifique uma fraude em uma declaração do IR, ela pode procurar outras declarações que contenham o mesmo endereço MAC, ou seja, que foram feitas em um mesmo computador e, provavelmente, por uma mesma pessoa. A partir daí a Receita pode verificar mais cuidadosamente essas outras declarações encontradas.

Ainda, a partir dos dados dos contribuintes a Receita pode obter a localização física (estado, cidade, rua número,…), mais ou menos precisa, dos endereços MAC. E então, pode até mesmo determinar quais (quantos) computadores estão em uma mesma residência, sala comercial, prédio, cidade ou estado. Além de ser possível outras deduções relacionadas às marcas dos equipamentos.

5 Conclusão

O endereço MAC pode revelar muito mais informações do que o endereço IP. Com ele é possível identificar de forma única, em termos mundiais, um computador ou outro equipamento qualquer que possua uma interface de rede. Essa informação não é uma informação fiscal e vem sendo capturada desde 2010 pela Receita Federal sem o consentimento e ciência do dono do

equipamento e do contribuinte. Vale destacar que ambos não são necessariamente a mesma pessoa. Não há nenhuma menção à coleta dessa informação seja no programa, política de privacidade ou qualquer outro documento disponibilizado pela Receita em seu site.

Assim como os endereços MAC são capturados, outras informações também poderiam ser facilmente capturadas e enviadas à Receita Federal como, por exemplo: listagem de arquivos de declaração (.DEC) existentes no disco do computador, nome de programas em execução, nome do usuário no sistema operacional, dados de licença do sistema operacional, etc.

Certamente há espaço aqui para um estudo técnico-jurídico, principalmente no que diz respeito ao anteprojeto da lei para proteção de dados pessoais. Não só no que diz respeito ao caso específico analisado mas à diversas outras iniciativas, principalmente privadas, que também fazem uso de dados coletados sem o devido consentimento dos usuários.

(*) *Vinícius da Silveira Serafim* é graduado em Ciência da Computação pela Universidade de Passo Fundo (1999) e Mestre em Ciência da Computação pela Universidade Federal do Rio Grande do Sul (2002). Trabalha como consultor em segurança da informação desde 2001 e atua como professor de cursos de extensão, graduação e pós-graduação desde 1999.

Anexo 1 - Função fa()

```
  public static String fa()
  {
    String str = "";
    Object localObject = fb();
    try
    {
      if ((localObject != null) && ((localObject =
NetworkInterface.getByInetAddress((InetAddress)localObject)) != null) && ((localObject =
((NetworkInterface)localObject).getHardwareAddress()) != null) && (localObject.length ==
6)) {
        str = String.format("%1$02x%2$02x%3$02x%4$02x%5$02x%6$02x", new Object[]
{ Byte.valueOf(localObject[0]), Byte.valueOf(localObject[1]),
Byte.valueOf(localObject[2]), Byte.valueOf(localObject[3]), Byte.valueOf(localObject[4]),
Byte.valueOf(localObject[5]) }).toUpperCase();
      }
      if (str.isEmpty())
      {
        Enumeration localEnumeration = NetworkInterface.getNetworkInterfaces();
        while (localEnumeration.hasMoreElements()) {
          if (((localObject = (NetworkInterface)localEnumeration.nextElement()) != null)
&& ((localObject = ((NetworkInterface)localObject).getHardwareAddress()) != null) &&
(localObject.length == 6))
          {
            str = String.format("%1$02x%2$02x%3$02x%4$02x%5$02x%6$02x", new Object[]
{ Byte.valueOf(localObject[0]), Byte.valueOf(localObject[1]),
Byte.valueOf(localObject[2]), Byte.valueOf(localObject[3]), Byte.valueOf(localObject[4]),
Byte.valueOf(localObject[5]) }).toUpperCase();
            break;
          }
        }
      }
    }
    catch (SocketException localSocketException)
    {
      (localObject = localSocketException).printStackTrace(System.err);
    }
    return str;
  }
```

Anexo 2 - Função fb()

```
  private static InetAddress fb()
  {
    String[] arrayOfString = { "www.receita.fazenda.gov.br", "www.google.com.br",
"www.uol.com.br", "www.globo.com", "www.terra.com.br", "www.estadao.com.br" };
    int i = 0;
    int j = 0;
    InetAddress localInetAddress2 = null;
    while ((i == 0) && (j < 6)) {
      try
      {
        InetAddress localInetAddress1 = InetAddress.getByName(arrayOfString[(j++)]);
        Object localObject1;
        int k = (localObject1 = new ServerSocket(0)).getLocalPort();
        ((ServerSocket)localObject1).close();
        k = k;
        (localObject1 = new DatagramSocket(k)).connect(localInetAddress1, k);
        localInetAddress2 = ((DatagramSocket)localObject1).getLocalAddress();
        i = 1;
      }
      catch (UnknownHostException localUnknownHostException)
      {
        (localObject2 = localUnknownHostException).printStackTrace(System.err);
      }
      catch (SocketException localSocketException)
      {
        (localObject2 = localSocketException).printStackTrace(System.err);
      }
      catch (IOException localIOException)
      {
        Object localObject2;
        (localObject2 = localIOException).printStackTrace(System.err);
      }
    }
    return localInetAddress2;
  }
```