Protótipo para exfiltração óptica de dados em máquinas fisicamente isoladas

Arthur Costa Lopes[1]*, **Diego F. Aranha**[1]

[1]Instituto de Computação (IC) – Universidade Estadual de Campinas (Unicamp)
arthur.lopes@students.ic.unicamp.com, dfaranha@ic.unicamp.br

Abstract. *We present a new approach for data exfiltration using a malicious storage device which subtly transmits data through blinking LEDs. This new approach may be used by an attacker trying to leak sensitive data stored in the device, such as credentials, cryptographic keys or a small classified document. An ideal application for this approach is when an attacker is capable of sneaking a malicious device inside a protected perimeter and has remote control over a camera inside such perimeter. We present several techniques for optimizing communication between transmitter and receiver, including application of error-correcting codes, achieving transmission rates up to 30 bits per second.*

Resumo. *Este trabalho apresenta uma nova abordagem para exfiltração de dados utilizando um dispositivo malicioso de armazenamento que transmite dados sutilmente utilizando LEDs infravermelho. A abordagem pode ser utilizada por um atacante para capturar dados armazenados no dispositivo, como credenciais ou chaves criptográficas. Um cenário ideal de aplicação ocorre quando o adversário é capaz de introduzir o dispositivo em um perímetro protegido e então coletar informação por meio de uma câmera sob seu controle remoto. Discutem-se técnicas para otimizar comunicação entre transmissor e receptor, incluindo a aplicação de códigos corretores de erro, obtendo taxas de transmissão de até 30 bits por segundo.*

1. Introdução

Com a ampla disponibilidade de sistemas de computação, organizações começaram a fazer uso extensivo de computadores para armazenar e trocar informações, muitas vezes de natureza extremamente sensível. A segurança da informação está se tornando cada vez mais essencial para proteger e impedir que esses dados vazem de diversas maneiras, como por exemplo, por atividades maliciosas de exfiltração. O conceito de exfiltração de dados consiste na extração de dados a partir de uma rede fechada depois de um *software* ou dispositivo malicioso se infiltrar em tal rede.

Nos últimos anos, as políticas de segurança começaram a considerar a ameaça da exfiltração de dados. Várias formas de prevenção têm sido recomendadas, incluindo o isolamento de uma máquina de qualquer rede externa (*air-gapping*) [Maas 2013]. Este trabalho tem como objetivo estudar novos métodos ópticos pouco intrusivos para exfiltração de dados, junto às suas limitações e eficiência. Ao antecipar o impacto desta ameaça, as organizações tornam-se capazes de proteger seus dados contra exfiltração, projetando contramedidas adequadas. Ao mesmo tempo, aprimorar os mecanismos de exfiltração é importante para informar ativistas e delatores sobre as formas plausíveis para recolher provas de comportamento desonesto e vazar informações de forma sutil e eficiente.

*Agradecimentos à FAPESP pela concessão da bolsa de iniciação científica, processo 2015/13876-7.

A abordagem proposta tem vantagens quando comparada com resultados recentes da literatura. Ao contrário de outras técnicas que exigem que um *software* malicioso seja instalado na máquina atacada para controle privilegiado do *hardware*, apenas um dispositivo de armazenamento portátil, *pendrive* ou HD externo, precisa ser inserido no ambiente alvo, possivelmente empregando técnicas de engenharia social [Tischer et al. 2016]. Depois que a unidade de armazenamento é ligada à máquina fisicamente isolada e os dados são recebidos, o dispositivo malicioso pode começar a transmitir dados capturados para o adversário, sem exigir um alto nível de privilégio ou interferir de qualquer forma com o sistema. Além disso, não há a necessidade do atacante recuperar o dispositivo mais tarde, uma vez que os dados são capturados por uma câmera ou por um *smartphone* no interior do ambiente. Dessa forma, não há intervenção física no sistema atacado, que poderia ser relevada posteriormente em uma investigação forense, em caso de suspeita. A exfiltração óptica de dados através de LEDs também oferece uma maior largura de banda em comparação com outros métodos, permitindo que o invasor obtenha dados de maneira mais eficiente e permitindo uma maior flexibilidade para ser usada em múltiplos cenários.

2. Trabalhos relacionados

Diversos trabalhos na área de exfiltração de dados foram realizados nos últimos anos utilizando diferentes abordagens, como LED do monitor [Sepetnitsky et al. 2014], som da ventoinha [Guri et al. 2016b], emissão de calor [Guri et al. 2015] e ondas de rádio do barramento USB [Guri et al. 2016a]. Tais abordagens são tipicamente demonstradas para máquinas sem conexão a redes externas, nas quais vazar dados é uma tarefa complexa. As velocidades de transmissão são comumente baixas e os mecanismos de recepção das informações são muitas vezes de difícil elaboração. Por exemplo, utilizar calor na exfiltração de dados impõe uma largura de banda extremamente limitada, devido ao tempo necessário para aquecer e esfriar o computador somente executando um componente de *software* em segundo plano. Outra desvantagem destes métodos é o elevado nível de intrusão, que leva à necessidade de modificar um periférico antes do mesmo ter sido usado no ambiente alvo, o que pode facilitar a detecção; ou que um *software* malicioso seja instalado na máquina atacada para um controle privilegiado do *hardware*. Do ponto de vista de desempenho, LEDs de um monitor impõem um limite de frequência de 25Hz e transmissão por calor enfrenta um limite ainda menor.

Embora haja propostas para exfiltração de dados usando dispositivos de armazenamento portátil [Zaddach et al. 2013, Clark et al. 2009], a abordagem óptica que buscamos é nova e não requer adulteração de *hardware* ou inserção de *malware* no dispositivo para que ele possa acessar e vazar dados confidenciais do computador. A Tabela 1 resume alguns aspectos de diferentes técnicas de exfiltração de dados encontradas na literatura com a proposta nesse projeto. O nível de intrusão alto se refere à necessidade de infecção com *software* privilegiado, enquanto níveis médio (dispositivo malicioso) e baixo (apenas programa em nível de usuário) relaxam esse requisito.

3. Protótipo para exfiltração óptica

No cenário considerado, assume-se primeiramente que o atacante é capaz de infectar uma máquina conectada (computador em rede ou *smartphone*) dentro do perímetro de segurança para controlar sua câmera. A câmera fica à espera que LEDs em seu campo de visão comecem a piscar de uma maneira específica, com uma sequência para sincronizar

Tipo	Velocidade	Intrusividade	Taxa de transmissão
LED do monitor (luz)	Alta	Alta	25Hz
Bitwhisper (calor)	Baixa	Baixa	1 a 8 bits/hora
Fansmitter (som)	Baixa	Alta	900 bits/hora
USBee (rádio)	Alta	Alta	20 a 80 bytes/segundo
Método apresentado	Alta	Médio	30 bits/segundo

Tabela 1. Comparação dos métodos de exfiltração estudados na literatura: luz [Sepetnitsky et al. 2014], som [Guri et al. 2016b], calor [Guri et al. 2015] e rádio [Guri et al. 2016a].

o transmissor e o receptor. Estes LEDs fazem parte de um dispositivo que funciona como uma unidade de armazenamento simples, embora tenha *firmware* modificado para transmitir dados específicos armazenados dentro da memória. O dispositivo previamente introduzido é então conectado a uma máquina isolada, da qual o atacante objetiva vazar informações. Após a câmera monitorar o LED e completar o protocolo de sincronização, os dados podem ser lidos e enviados para um servidor controlado pelo atacante. O projeto está dividido em duas partes: o algoritmo para transmitir dados através do meio óptico a partir do dispositivo de armazenamento e o sistema de captura e decodificação.

3.1. Transmissor

A plataforma escolhida foi uma placa Teensy2, compatível com Arduino. O *framework* LUFA[1] é utilizado para construir o *pendrive* com um cartão SD comum[2], que funciona como memória do dispositivo. Alguns LEDs finalizam o projeto.

A Figura 1 apresenta a versão atual do protótipo do transmissor. O protótipo é ligado a um computador através de uma porta USB comum e pode ser usado como um dispositivo de armazenamento regular, utilizando um protocolo de comunicação padronizado. A primeira versão do protótipo era equipada com um único LED, representando um indicador de atividade. Como a largura de banda ficou limitada, a solução adotada inclui múltiplos LEDs, ampliando substancialmente a capacidade de transmissão. Para evitar aparência suspeita, os LEDs emitem luz infravermelha, invisível ao olho humano. Além disso o dispositivo foi colocado em uma capa transparente, para parecer com um dispositivo real. As extensões ou nomes dos arquivos de interesse são colocados na memória do dispositivo antes de sua entrada no perímetro de segurança. Quando o dispositivo for utililizado, buscará os dados escolhidos e logo em seguida a transmissão será iniciada.

Figura 1. Protótipo do dispositivo.

3.2. Receptor

O receptor é representado por uma *webcam* comum, capturando 30 quadros por segundo na resolução de 720p. Para a parte de *software*, utilizamos a biblioteca gráfica

[1] `http://www.fourwalledcubicle.com/LUFA.php`
[2] `http://elasticsheep.com/2010/04/teensy2-usb-mass-storage-with-an-sd-card/`

OpenCV 3.0[3] para manipular as imagens da câmera e obter a informação transmitida pelos LEDs do transmissor. O componente de *software* identifica a fonte dos dados a serem transmitidos, mantendo o controle de cada LED e de seu comportamento. Após a sincronização, o *software* começa a decodificar a informação através de um código de correção de erros para detectar falhas na transmissão. A cadeia de bits corrigida pode ser armazenada na máquina remota ou pode ser enviada para um servidor web através da conexão móvel do *smartphone* ou da rede externa disponível no computador acoplado na *webcam*. Este servidor é controlado pelo atacante, proporcionando um refúgio seguro para os dados finalmente extraídos.

O algoritmo para a detecção de um LED é simples e se baseia na diferença entre dois quadros consecutivos para determinar quais as partes da imagem têm a mudança do estado do LED e se há ruído de fundo. Após o protocolo inicial terminar, o receptor sabe a localização de cada LED e pode monitorar os estados para receber mensagens. Os bits são transmitidos da maneira mais simples possível: o LED ligado equivale a um valor 1, e um valor 0 caso contrário.

3.3. Protocolo

O protocolo de transmissão começa a partir da sequência específica de bits transmitidos no início da mensagem. A sequência é usada para marcar quando a mensagem está começando novamente para que o receptor possa parar de decodificar novos bits e começar a procurar pedaços de informações que foram perdidas durante as mensagens anteriores. Outra finalidade é dar à câmera algum tempo para encontrar os LEDs na primeira leitura.

Após o protocolo inicial ser concluído, os bits da mensagem são transmitidos através dos LEDs. No entanto, este tipo de transmissão ruidosa introduz alguns erros nos dados que estão sendo enviados. Por causa do tempo que o LED leva para acender ou apagar, um bit às vezes é lido incorretamente, geralmente no final de uma sequência de bits iguais. A maneira mais fácil encontrada para lidar com este problema foi utilizar o código de Hamming que permite a correção de um bit por palavra. Esta escolha foi suficiente para acabar com os erros nos testes e tem resolvido também os casos em que o erro ocorre quando o LED começa a apagar e a câmera captura esse momento transitório, resultando em alguns bits trocados.

Uma característica importante sobre o protocolo escolhido foi a baixa quantidade de bits de paridade, não afetando a largura de banda significativamente. Tipos mais complexos de códigos de correção de erros foram consideradas, mas descartadas devido a um número muito maior de bits de paridade.

4. Resultados preliminares

O principal objetivo é transmitir o conteúdo do cartão SD, logo o dispositivo consegue escolher um arquivo específico da memória para enviar. Estes dados são transmitidos através de múltiplos LEDs, cada um representando um bit diferente. Em termos de velocidade, o protótipo é capaz de transmitir 10-15 bits por segundo por LED, sendo a principal limitação encontrada no tempo necessário para acender ou desligar o LED completamente. Como o protótipo faz uso de 2 LEDs diferentes simultâneos, velocidades de transmissão até 30 bits por segundo foram observadas.

[3] `http://opencv.org/about.html`

O método já pode ser utilizado para enviar dados pequenos como senhas e chaves criptográficas, devido à sua relativamente baixa velocidade quando comparada ao tamanho típico de arquivos, apesar de ser rápida o suficiente para transmitir essas informações em poucos segundos. Documentos pequenos também podem ser transmitidos se o atacante tiver tempo suficiente disponível. Um desafio permanente do projeto é conseguir a máxima largura de banda possível para maximizar a eficácia da exfiltração. A maior limitação é a taxa de quadros da câmera na recepção, porque cada LED é capaz de transmitir mais de 30 bits por segundo. Essa foi a principal razão que levou a usar vários LEDs simultâneos. Esta modificação permitiu que a taxa de quadros fosse muito mais elevada, com apenas um pouco mais de esforço. Os resultados preliminares não estão tão longe de alguns trabalhos relacionados publicados [Sepetnitsky et al. 2014]. Dadas as limitações da abordagem óptica, a velocidade atingida parece promissora, tornando possível enviar dados sensíveis em apenas alguns segundos, e alguns pequenos arquivos em minutos.

5. Conclusão

Com os resultados obtidos pretende-se continuar o projeto aprimorando o protótipo para torná-lo apto à uma aplicação real. O principal objetivo é aumentar a velocidade de transmissão para possibilitar a transferência de arquivos de forma discreta e eficaz. Como o intuito do projeto é estudar formas novas de exfiltração, o estudo das formas de prevenção para tal método se mostra muito importante, para evitar de forma eficiente que dados sensíveis sejam vazados.

Referências

Clark, J., Leblanc, S., and Knight, S. (2009). Hardware trojan horse device based on unintended USB channels. In *NSS*, pages 1–8.

Guri, M., Monitz, M., and Elovici, Y. (2016a). Usbee: Air-gap covert-channel via electromagnetic emission from USB. *CoRR*, abs/1608.08397.

Guri, M., Monitz, M., Mirski, Y., and Elovici, Y. (2015). Bitwhisper: Covert signaling channel between air-gapped computers using thermal manipulations. In *IEEE CSF*, pages 276–289.

Guri, M., Solewicz, Y. A., Daidakulov, A., and Elovici, Y. (2016b). Fansmitter: Acoustic data exfiltration from (speakerless) air-gapped computers. *CoRR*, abs/1606.05915.

Maas, P. (2013). How Laura Poitras helped Snowden spill his secrets. New York Times: `http://www.nytimes.com/2013/08/18/magazine/laura-poitras-snowden.html`.

Sepetnitsky, V., Guri, M., and Elovici, Y. (2014). Exfiltration of information from air-gapped machines using monitor's LED indicator. In *IEE JISIC*, pages 264–267.

Tischer, M., Durumeric, Z., Foster, S., Duan, S., Mori, A., Bursztein, E., and Bailey, M. (2016). Users Really Do Plug in USB Drives They Find. In *IEEE Security & Privacy*, San Jose, California, USA.

Zaddach, J., Kurmus, A., Balzarotti, D., Blass, E., Francillon, A., Goodspeed, T., Gupta, M., and Koltsidas, I. (2013). Implementation and implications of a stealth hard-drive backdoor. In *ACSAC*, pages 279–288. ACM.

