Privacy Negotiation for TLS - Selectable SNI or SNO: Server Name Omission
TCP Increased Security (tcpinc) Working Group
Berlin, July 19, 2016

Dave Plonka <plonka@akamai.com>
Outline

• **Premise**: SNI leaks what could be considered private information.

• **Privacy Challenge**: Rendezvous-based Traffic Classification

• **Proposal**: selective Server Name Omission
SNI Leaks Private Information

• **Premise**: TLS with SNI leaks what could be private information and makes traffic classification much easier, sometimes trivial.
SNI Leaks Private Information

• **Premise**: TLS with SNI leaks what could be private information and makes traffic classification much easier, sometimes trivial.

• SNI was introduced c. 2004, currently specified by RFC 6066, “to facilitate secure connections to servers that host multiple 'virtual' servers at a single underlying network address.”
SNI Leaks Private Information

- **Premise**: TLS with SNI leaks what could be private information and makes traffic classification much easier, sometimes trivial.
  - SNI was introduced c. 2004, currently specified by RFC 6066, “to facilitate secure connections to servers that host multiple 'virtual' servers at a single underlying network address.”
  - Unfortunately, for applications that use it, SNI is “always on,” *i.e.*, sent unconditionally.
    - Presumably this was to avoid a round-trip-time to negotiate its inclusion during TLS setup.
SNI Leaks Private Information

• **Premise**: TLS with SNI leaks what could be private information and makes traffic classification much easier, sometimes trivial.
  • SNI was introduced c. 2004, currently specified by RFC 6066, “to facilitate secure connections to servers that host multiple 'virtual' servers at a single underlying network address.”
  • Unfortunately, for applications that use it, SNI is “always on,” *i.e.*, sent unconditionally.
    • Presumably this was to avoid a round-trip-time to negotiate its inclusion during TLS setup.
  • Virtual hosting and, therefore, SNI are *unnecessary* with IPv6; servers typically have $2^{64}$ addresses available.
Rendezvous-based Traffic Classification

- **Rendezvous-based Traffic Classification:** using DPI on Rendezvous traffic (e.g., unencrypted DNS and SNI) with transport information to flexibly classify traffic that has been passive observed.
  - Developed as flexible way to classify traffic in real-time at high-volume, with little DPI, and as a way to classify encrypted traffic.

- SNI is a TLS **rendezvous mechanism** that selects the server-side peer by name using clear-text information that is available by DPI at low-volume.
  - This has been used both as a basis for classification and ground-truth to validate and improve classifiers.
Rendezvous-based Traffic Classification

  - Treetop (Plonka & Barford, 2008-2013)
  - DN-Hunter / tstat (Mellia *et al.*, 2012-2016)
  - DNS-Class (Foremski *et al.*, 2014)

- **Patents:**
  - Apparatus and method for classifying network packet data (US7907543, 2011)
  - Discerning web content and services based on real-time DNS tagging (US8819227, 2014)
Rendezvous-based Traffic Classification: 2016

• “[By leveraging hostname to address associations ...] Our results show that up to 55% of web traffic can be identified relying solely on addresses.” (Trevisan et al., 2016)
SNO: selective Server Name Omission

- **Proposal:** Selectively omit or obscure Server Name Indication (SNI)
SNO: selective Server Name Omission

- **Proposal**: Selectively omit or obscure Server Name Indication (SNI)

- TCP-ENO is a way to negotiate increased privacy and, thus, seems a candidate method by which a server could suggest clear-text SNI preamble should be omitted, *i.e.*, “Turn privacy up to 11.”
SNO: selective Server Name Omission

- **Proposal:** Selectively omit or obscure Server Name Indication (SNI)

- TCP-ENO is a way to negotiate increased privacy and, thus, seems a candidate method by which a server could suggest clear-text SNI preamble should be omitted, *i.e.*, “Turn privacy up to 11.”

- Likely would work in concert with DPRIVE (RFC7858) and DANE as it, ultimately, wants the server not to expose the service name in clear-text, as with the certificate.
SNO: selective Server Name Omission

• **Initial feedback includes:**
  - “My main fear is delaying TCP-ENO further.”
  - “Perhaps finishing up now with the tiny set of codepoints already considered is right if the WG could add other ones later.”
  - “I think it’d still be good to get folks’ reactions to this idea now.”

• **Technical issues:**
  - Does it affect downgrade attacks by (active) man-in-the-middle?
SNO: selective Server Name Omission

- **Position:** Omitting clear-text SNI when accessing TLS-based services is a key ingredient in some recipes for a more private Web and Internet.

- **Where and when:** Is TCPINC the place for this work? Why or why not?
Privacy Negotiation for TLS - Selectable SNI or SNO: Server Name Omission

Thanks!
Questions, Comments?