Counteract traffic analysis attacks on VoIP media flows

Ge Zhang
Karlstad University

Nowadays, Voice over Internet Protocol (VoIP) which enables voice conversation remotely over packet switched networks gains much attention for its low costs and flexible services. However, VoIP calling anonymity, particularly to withhold “who called whom”, is difficult to achieve since VoIP infrastructures are usually deployed in an open networking environment (e.g., the Internet). Even if the VoIP packet payloads are encrypted, wiretapping attackers in the communication channels can still get “who called whom” from the packet-headers.

Nowadays, many VoIP service providers relay media flows for users. The original purpose for doing so is to enable traffic traverse firewalls or NATs. However, a good side-effect is that the wiretapping attackers cannot directly profile the calling records from observed packet-headers as the flows are not end-to-end built.

Nevertheless, the relay solution cannot counteract sophisticated attackers who can link two corresponding subflows on both side of the relay. Actually, there are many proposed methods to achieve this: (1) Attackers can introduce timing watermark in an ingress flow by delaying selected packets. The timing watermarks usually cannot be detected and removed by the relay. Thus, the attacker can find the corresponding egress flow by decoding the watermark on the other side. (2) Another method takes advantage of human conversation behaviors: When one is speaking, another usually is silent. Thus, attackers can pair flows on two sides by detecting the silent and speaking periods from the flows.

Taking these two attacks into account, we propose the “defensive dropping” method in VoIP: A VoIP user-agent sends media packets to the relay in a constant rate with the packets during periods of silence are marked. Then, the relay drops some silence packets and forwards the remaining ones to their destinations. Since silence packets are less meaningful, dropping them will not impact the quality of conversations too much. The purpose of dropping silence packets is to remove watermarks in a certain degree. Thus, more packets dropped, more difficult to recover the watermarks. Unfortunately, however, dropping too much silence packets could expose the silent and speaking periods in the flows, yet leading to the second attack.

The result of our experiments shows that the dropping rate must be carefully selected (around 10% in our examples).