AVOID: Automatic Verification Of Internet Data-paths

Alexander Marder[1], Jon Larrea[2], kc claffy[3], Erik Kline[4], Kyle Jamieson[5],
Bradley Huffaker[3], Lincoln Thurlow[4], and Matthew Luckie[3]
[1]Johns Hopkins University, [2]Revelare Networks, [3]UCSD CAIDA, [4]USC/ISI, [5]Princeton University

I. Introduction

Department of Defense (DOD) use of commercial networks entails unprecedented reliance on untrusted third-party communications infrastructure, and the associated risk of exposing DOD communications to an adversary. Traversing adversary-controlled infrastructure allows DOD's adversaries to recognize, disrupt, or extract intelligence even from encrypted communications. The resulting arms race of obfuscation vs intelligence techniques is inherently limited: with each new obfuscation, DOD can never know if it fools the adversary, or if the adversary is lulling DOD into a false sense of security.

We believe the next great capability leap for operating through commercial networks will likely come from sophisticated analytics that provide situational awareness of the threats within the communications infrastructure, and implementations that dynamically route communications along benign paths. These systems will restructure communication paths to avoid adversary-controlled infrastructure in cellular and Internet networks, complementing existing DOD defenses and keeping communications unobservable by the adversary.

Our demonstration will show results of our vision for this emerging conceptual framework, that we call AVOID. We will demonstrate a phone app that can identify potentially malicious base stations in cellular networks, and automatically connect to benign base stations. We will also demonstrate that the same phone—combined with a topology-aware overlay—can avoid adversary-controlled infrastructure in ISP networks outside the wireless network.

II. AVOID Base Station Vendors

In cellular networks, the radio access network (RAN) is responsible for providing wireless access to users. The main components of the RAN are the base stations that mediate access to the shared wireless channel. The base stations are the entry into the wireless cellular network, bridging between user equipment (UEs) and the core of the wireless network.

Adversaries can compromise DOD communications in the RAN in two ways. One way is exploiting intentionally placed backdoors in 4G and 5G base station equipment that provide adversaries with the ability to monitor, modify, and exfiltrate traffic for offline analysis and future operations [1]. Evidence suggests that the Chinese base station vendors Huawei and ZTE—comprising 1/3 of the global base station market—have compromised supply chains that include potential backdoor access [2]. A second way is deploying surveillance base stations that trick UEs into connecting in order to tie service members to devices, learn patterns of life, or locate users. A sophisticated surveillance base station could even authenticate subscribers, allowing it to capture data from the UEs.

A. Proposed Approach: Classify Vendors Based on Behavior

3GPP specifies much of the base station behavior and protocols to ensure interoperability, but it remains possible to classify base station vendors based on how they are configured and act. Design decisions and proprietary algorithms differ across base station vendors, and UEs can observe differences in the Radio Resource Control (RRC) configurations and Medium Access Control (MAC) layer. By creating classifiers that map these differences to vendors, UEs can classify vendors with high confidence and detect when surveillance equipment deviates from expected behavior.

The vendor classifiers could help advanced reconnaissance teams assess risks posed by the RAN before deploying personnel or equipment in an area. They could also help detect custom private deployments by adversary forces, since those base stations will likely appear different in the control channel than carrier equipment. The most ambitious use of vendor identification—which we will demonstrate—is to automatically avoid malicious vendors and surveillance equipment by placing intelligence on UEs.

B. AVOID-Vendor Demonstration: COTS Android Phone

We created an initial prototype of the proposed RAN solution—*AVOID-Vendor*—that secures individual commercial off-the-shelf (COTS) Android devices, that can protect communications for service members and special operators overseas (Fig. 1). *AVOID-Vendor* leverages RRC and MAC information extracted directly from the cellular modem within the phone, and uses it to classify base station vendors in the RAN. When run on phones with multiple SIM cards, *AVOID-Vendor* uses modem manipulation to ensure the phone only connects to benign base stations.

The AVOID app running on the phone orchestrates the system and implements the *AVOID-Vendor* strategy. Every time the phone connects to a new base station, AVOID immediately classifies the base station with three possible actions. (1) AVOID takes *no action* when the classifier returns a high confidence prediction that the base station is benign. (2) If the classification results in a high confidence vendor

Funded in part by NSF ITE-2326928.

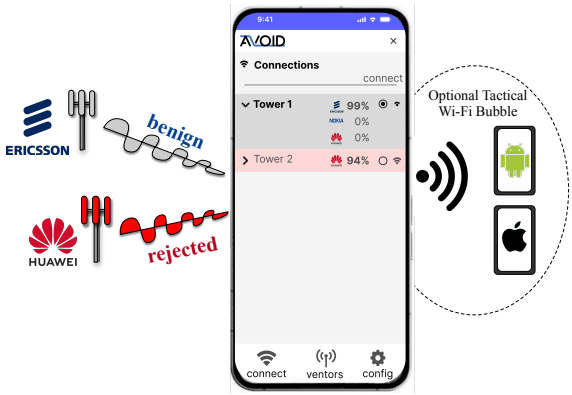

Fig. 1. AVOID app automatically rejects connecting to the Huawei base station, and instead forces the phone to connect through the benign Ericsson base station. The AVOID app can also optionally create a tactical Wi-Fi bubble that secures cellular communcations for other nearby devices.

prediction of Huawei or ZTE, then AVOID *switches carriers*. (3) Finally, low confidence vendor classification suggests that the base station vendor is not known to the classifier, and therefore presents a potential risk to communications. AVOID instructs the modem controller to *blacklist* the base station to avoid connecting to surveillance equipment.

We will demonstrate that the AVOID app prevents a phone from connecting to specific base station vendors in commercial U.S. cellular networks. In Washington DC, AT&T, Verizon, and T-Mobile all use different base station vendors, and we will configure AVOID to treat one or more of those vendors as malicious. AVOID will identify the base stations from that vendor, and switch carriers to connect to base stations made by a different vendor. Our demonstration will also include two optional external modules – an SDR module that passively classifies base stations, and a Wi-Fi module that secures communications for nearby devices.

III. AVOID Malicious ISP Infrastructure

Transmitting information through the untrusted public Internet can expose DOD communications to the sophisticated traffic recognition and disruption capabilities that our adversaries possess. As soon as DOD transmits a packet into a commercial wireless network destined for the public Internet, DOD loses control over how that packet reaches the destination. Any Internet Service Provider (ISP) headquartered in an adversarial nation-state, any ISP infrastructure residing in adversary-controlled territory, and potentially even routers manufactured by certain vendors, can be compelled to subject traffic to the adversary's intelligence regime. These capabilities are advanced and complex, using passive and active techniques to thwart attempts at traffic obfuscation [3], including VPNs and Tor's pluggable transports.

A. Proposed Approach: Underlay-Aware Overlay Routing

DOD needs the communications infrastructure equivalent to clean supply chains. UEs have no control over the router-level Internet paths that transmitted data takes and advances

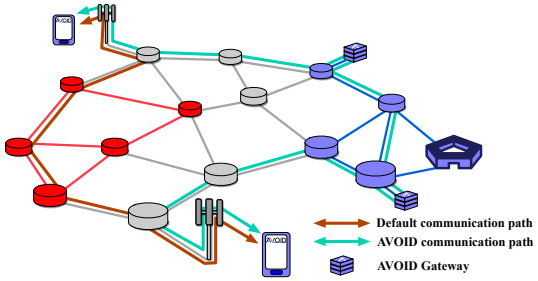

Fig. 2. The AVOID system will recognize adversary-controlled infrastructure (red) and route communications along benign paths to the AVOID gateways in the DOD network (blue).

in obfuscation and VPN architectures cannot fully overcome the fundamental challenge of an untrustworthy underlying infrastructure substrate. DOD needs overlay technology that is aware of the underlying network topology, including the geopolitical and economic attributes of that topology. The Internet's highly distributed control prevents per-router-hop control of traffic crossing independent commercial networks. But one essential feature of global IP networks—the pervasive use of destination- or flow-based forwarding—means that with sufficiently rich understanding of the underlying topology, one can use strategic selection of source and destination IP addresses to indirectly choose an end-to-end path.

B. AVOID-Path Demonstration: Avoid Adversary Location

We implemented an initial prototype of a topology-aware overlay, called *AVOID-Path* (Fig. 2). On initial connection, the AVOID app on a COTS Android phone reaches out to a bootstrapping server that an initial overlay gateway node. These gateway nodes are implemented as Docker containers and can be placed anywhere there is a secure path to the rest of the DOD network; e.g., forwarding operating bases. We include analytics that determine the location, operating network, and vendor for routers in the communication paths. If the analytics indicate that the path to the initial gateway includes risky infrastructure, the AVOID app will connect to another gateway reachable over benign paths.

Our demonstration will show the UE attempt a connection to a server without the communication traversing any router within a specified set of US states. We will represent the path visually on a map, and show the results of applying the geolocation, network operator, and router vendor analytics over the path. Then the UE will join the overlay and connect to an initial gateway, and we will map the path. This path will traverse problematic locations as well, requiring the UE to switch to another gateway. Finally, we will show that the analytics accept the new path to the gateway.

References

[1] B. Pancevski, "U.S. Officials Say Huawei Can Covertly Access Telecom Networks," Washington Post, 2020.
[2] C. Reichert, "U.S. finds Huawei has backdoor access to mobile networks globally, report says," CNET.
[3] J. Beznazwy and A. Houmansadr, "How China Detects and Blocks Shadowsocks," in *ACM IMC*, 2020.