ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/317018072

Modeling inter-signal arrival times for accurate detection of CAN bus signal injection attacks: a data-driven approach to in-vehicle intrusion detection

Conference Paper · April 2017

DOI: 10.1145/3064814.3064816

CITATIONS
23

READS
124

5 authors, including:

Michael R Moore
Oak Ridge National Laboratory
50 PUBLICATIONS **158** CITATIONS
SEE PROFILE

Robert Bridges
Oak Ridge National Laboratory
35 PUBLICATIONS **178** CITATIONS
SEE PROFILE

Frank L. Combs
Oak Ridge National Laboratory
5 PUBLICATIONS **24** CITATIONS
SEE PROFILE

Stacy J. Prowell
Oak Ridge National Laboratory
52 PUBLICATIONS **533** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

dissertation View project

Annual security workshop at Oak Ridge National Laboratory View project

All content following this page was uploaded by Robert Bridges on 05 January 2018.

The user has requested enhancement of the downloaded file.

Modeling inter-signal arrival times for accurate detection of CAN bus signal injection attacks*

A data-driven approach to in-vehicle intrusion detection

Michael R. Moore,[1] Robert A. Bridges,[2] Frank L. Combs,[3] Michael S. Starr,[1] & Stacy J. Prowell[2]
[1]Global Security Directorate, [2]Computer & Computational Sciences Directorate, [3]Energy & Environmental Sciences Directorate, Oak Ridge National Laboratory
1 Bethel Valley Road, Oak Ridge, TN 37831, USA
[mooremr,bridgesra,combsfl,starrms,prowellsj][@ornl.gov]

ABSTRACT

Modern vehicles rely on hundreds of on-board electronic control units (ECUs) communicating over in-vehicle networks. As external interfaces to the car control networks (such as the on-board diagnostic (OBD) port, auxiliary media ports, etc.) become common, and vehicle-to-vehicle / vehicle-to-infrastructure technology is in the near future, the attack surface for vehicles grows, exposing control networks to potentially life-critical attacks. This paper addresses the need for securing the controller area network (CAN) bus by detecting anomalous traffic patterns via unusual refresh rates of certain commands. While previous works have identified signal frequency as an important feature for CAN bus intrusion detection, this paper provides the first such algorithm with experiments using three attacks in five (total) scenarios. Our data-driven anomaly detection algorithm requires only five seconds of training time (on normal data) and achieves true positive / false discovery rates of 0.9998/0.00298, respectively (micro-averaged across the five experimental tests).

CCS CONCEPTS

•Security and privacy →Artificial immune systems;

KEYWORDS

CAN bus, in-vehicle security, anomaly detection, signal injection

ACM Reference format:
Michael R. Moore,[1] Robert A. Bridges,[2] Frank L. Combs,[3] Michael S. Starr,[1] & Stacy J. Prowell[2]. 2017. Modeling inter-signal arrival times for accurate detection of CAN bus signal injection attacks. In *Proceedings of Cyber & Information Security Research Conference, Oak Ridge, TN, USA, April 04 - 06, 2017 (CISRC '17),* 4 pages.
DOI: http://dx.doi.org/10.1145/3064814.3064816

*This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan http://energy.gov/downloads/doe-public-access-plan.
Parts of this research performed at the Vehicle Security Lab at the National Transportation Research Center.
[1]Author Michael S. Starr, Lt Col, USAF Fellow

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of the United States government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.
CISRC '17, Oak Ridge, TN, USA
© 2017 Copyright held by the owner/author(s). Publication rights licensed to ACM.
978-1-4503-4855-3/17/04…$15.00
DOI: http://dx.doi.org/10.1145/3064814.3064816

1 INTRODUCTION

A modern vehicle relies on scores of engine control units (ECUs), which are embedded computers controlling the vehicle's many subsystems. Because of the number of ECUs, dedicated connections for all ECU traffic is unfeasible and a single bus allowing all signals to be broadcast to all ECUs is standard. In particular, we focus on the high-speed (125Kbs-1Mbs) controller area network (CAN) bus used for much of modern vehicle communications. Because ECUs control most of the vehicle's functions (sensors, lights, braking, etc.), it follows that adversarial manipulation of signals on the CAN bus has potentially severe consequences. Exacerbating the potential for interference is the proliferation of external connections with the vehicle control network, including USB ports, WiFi, Bluetooth, and the mandatory on-board diagnostic (OBD-II) port that gives direct access to vehicle buses. Near-future advancements, including vehicle-to-vehicle and vehicle-to-infrastructure wireless communication, increase the need for vehicle network security. Protecting these critical control networks has led to increasing study of their vulnerabilities and mitigations for those vulnerabilities. [2, 9, 10]

CAN bus signals are indexed by a process ID (PID), specified in the packet header, and are generally associated with a fixed function (running lights, sensors, door locks, etc.) The specific PID-to-function mapping of signals is dependent on the make and model; e.g., signals with PID 3A1 may code for the brake lights in one make/model, but something different in another. This mapping poses problems for creating *universally* effective offensive and defensive cyber capabilities.

This work relies on the observation that most PID signals are sent regularly and redundantly. Command injection attackers for these PIDs' functions, therefore, need to produce regular, redundant signal injections to achieve a desired response in the vehicle's actions, and we define this class of attacks as *regular-frequency signal injection attacks.* Our hypothesis is that by modeling and detecting anomalies in the inter-signal wait times we can exploit the regularity of the CAN bus signals and can produce an accurate detection capability for this well-defined class of attacks.

To test our proposed detector, we define and execute three signal injection attacks. This serves to illuminate both the ease of execution and the potential for danger of these attacks. We present accuracy results of the detector under both normal (non-attack) and attack conditions. Rather than disclose details of vulnerabilities exploited, we have informed the vendor. For this reason the make, model, and production year of the testing car, as well as the injected PIDs and values, are not included.

Our detector does not require labeled attack data, and only requires a few seconds of normal (non-attack) CAN bus observations to train. Unlike signature-based detection methods, our system promises defense against novel attacks; more explicitly, our system is designed to detect any attack that disrupts the frequency of the CAN bus packets. Our initial empirical results support this goal with near perfect true positive and false detection rates. Because it depends only on the regularity of the inter-signal arrival times, the detection capability eliminates the need to reverse engineer the signal-to-effect mapping of the CAN signals for each make and model. Although outside the scope of this paper, we have begun engineering work to create a prototype (aftermarket) detection technology using the OBD-II plugin running our algorithm and illuminating lights to indicate anomalies in CAN signals' frequency.

Related Work

Our implementation of the signal injection attacks used an Arduino board connected to the vehicle via the OBD-II port, allowing remote access via bluetooth. Previous implementations of signal injection attacks both wireless and otherwise (similar to ours) are found throughout the literature [2–4]. Vehicle network intrusion detection, particularly, via anomaly detection, is an emerging area of research [3, 5, 6, 8]. Both Hoppe et al. [3] and and Müter et al. [7] identify CAN bus signal frequency as an important feature for signal injection detection, but no algorithm, detection experiments, or accuracy results are presented. This work presents the first detailed signal-frequency-based CAN-bus anomaly detection algorithm with experimental results. Broster et al. [1] develop a general system for detecting babbling node failures on event-triggered systems (in particular CAN) by knowing a priori timing constraints between signal triggers and infer if the constraints are broken from observed signal times. While our method leverages timing of signals, it requires no a priori system-specific knowledge and targets a more general class of attacks but is CAN specific.

2 REGULARITY OF SIGNAL FREQUENCY

Analyzing the CAN bus signals for normal settings (50s—25s with car on and engine off, plus 25s with car on and engine on), we make a few critical observations. First, nearly all (110 of 120 observed) PIDs' signals are *regularly occurring*, meaning that for a fixed id a signal is sent repeatedly at a fixed rate with little noise. As a specific example, the runtime lights signal is sent at a fixed interval with one of two values, indicating on/off; hence, even when not in use, the lights singal is sent. Explicitly, our observation on this regularity is that the time between signals (referred to as inter-signal arrival time or wait time) never differs more than a 24% from expectation, where expectation is computed as the mean of the observations. See Figure 1. Additionally, Figure 1 gives a histogram of the 117 (of 120) PIDs whose signals occur at a frequency greater than 1hz. The takeaway from these two histograms is that the vast majority of PID signals occur regularly, with nearly fixed frequency, and at a relatively high rate.

Following our observations during testing, we assume the value of a function remains constant until receipt of a signal with a new value. Consequently, *an adversary that wishes to control a subsystem via CAN bus signal injection must send the appropriate signals (PID and value) at a rate at least as fast as that process's ambient signal rate to sustain the desired value.* With this motivation, we define a "regular-frequency signal injection attack" as a repeated injection of a CAN bus signal with fixed PID and at a rate greater than or equal to the expected rate of that signal.

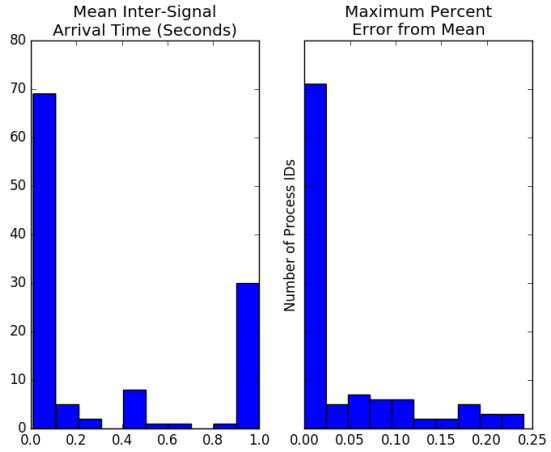

Figure 1: Left histogram displays the number of signals by mean inter-signal arrival times rounded to the nearest 0.001s. Note that three of 120 signals have mean inter-signal arrival time greater than 1.0005s. and are not depicted. (Right histogram) For each PID, the maximum percent error of an observed inter-signal wait time from the mean wait time is computed. Of 120 PIDs, 110 exhibited all signals within 24% of the expected wait time. Right histogram present these "regularly occurring" signals binned by the maximum observed percent error.

3 HYPOTHESIS, DETECTOR, & EXPERIMENT

In light of the regularity of signal frequencies, we hypothesize that a simple anomaly detection system monitoring the inter-signal wait times of CAN bus traffic will provide accurate detection of a regular-frequency signal injection attacks. As a proof of concept we present an experiment using three attacks during two normal vehicle operation modes, engine on/off. The first (respectively, second) attack turns runtime lights (respectively, reverse lights) from on (normal) to off (attack). These two attacks operate by repeatedly injecting the desired signal into the CAN bus at a rate high enough to suppress the ambient signal also sent regularly by the car; hence these are regular-frequency signal injection attacks by necessity. These two attacks are designed to be covert, changing the car's actions without the driver knowing, and their success was easily empirically verifiable. The third attack is a (less stealthy) denial-of-service (DOS) attack we discovered, which upon repeated signal injection of a particular signal, forces the car to turn off. See caption in Figure 3 for more description. Our experimentation reveals that the successful DOS (the car shutting off) by this particular signal also requires a regular-frequency signal injection. Finally, we note that our hardware implementation allowed blue-tooth connection

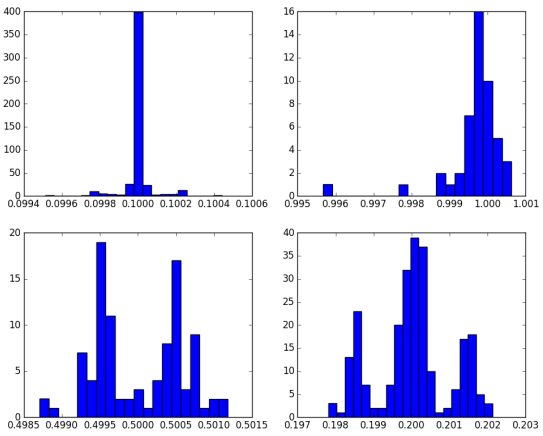

Figure 2: Disparity of the distributions of inter-signal arrival times depicted. (From top left to bottom right) PIDs exhibit extremely peaked unimodal, skewed unimodal, bimodal, and trimodal wait-time distributions.

to the Arduino board and, therefore, is triggerable from outside the vehicle. The simplicity of the hardware and the attack illustrates the ease with which a car can be compromised (provided an adversary could at some point access the OBD-II port near the steering column). In particular it demonstrates the urgent need for intrusion detection and prevention capabilities in this space.

3.1 Modelling Inter-Signal Arrival Times

It is natural to model each PID's signal occurrence times as samples of a stationary, Markov process—i.e., the time of a signal only depends on the time of the previous signal, and this wait time follows a fixed distribution. To do so, we estimate the wait-time distribution from observations and find that these distributions are very dissimilar across processes. See figure 2. In particular, this disparity dissuades us from simply fitting a Gaussian and identifying wait times with low z-score. Rather we exploit the relative regularity of the signals by using the percent of absolute error of an observed wait time from expectation.

Explicitly, we build an anomaly detector for each PID's signal stream. To train the models, we simply observe a few seconds of CAN data, and for each PID record the times that processes' signals occur. Denote the signal times for a given PID $t_0, t_1, ..., t_n$, so we have observed n wait times, namely, $\Delta_1 := t_1 - t_0, ..., \Delta_n := t_n - t_{n-1}$. From the $\{\Delta_i\}_i$ we find the maximum observed error from expectation, $m := \max_i |\Delta_i - \mu|$ where $\mu := \sum_i \Delta_i / n$, and set the threshold for absolute error from expectation as m plus 15% of μ. That is, future observations are flagged if the time between signals differs in absolute value from the expected wait time by more than $\alpha := m + 0.15 * \mu$. By design if two ECUs send CAN signals at the same time, the signal with lower PID (3-digit hex) takes precedence; hence, a dropped signal and, therefore, doubled wait time will occur on occasion. To accommodate this feature and limit false positives, we require three flagged wait times consecutively to issue an alert. Altogether there are 120 models (one for each PID); each will flag unusually short/long inter-signal times and produce an alert upon three consecutive anomalies. For successful results, our method depends on the few seconds of training observations being normal, i.e., with no attacks present.

3.2 Experiment & Results

Here we present our first tests of the above anomaly detector. Each data set includes one of the three aforementioned regular-frequency signal injection attacks occurring for about 29-45s and surrounded by normal (non-attack) data. Figure 3 gives details of the five test sets and results. For each data set, only the first five seconds of normal data is used to train the models, then anomaly detection as described above is run. Recall that each CAN bus packet is flagged as anomalous if and only if the wait time between it and the last signal with that PID is too long/short, and an alert is raised if and only if a packet is the third consecutive anomalous packet with that PID. To display the results in Figure 3 we count and plot the number of alerts in each .025s. non-overlapping time window. Micro-averaged rates are as follows: true positive rate = 0.9998, false positive rate = 0.00294, false detection rate = 0.00298.[1]

4 CONCLUSIONS & FUTURE WORK

Our initial investigations of CAN bus traffic have concluded that conditioned on PIDs, the signals are exceptionally regular in frequency while still exhibiting a wide variety of wait-time distributions. Most notably, we have two identified notable consequences of this regularity, namely, (1) that many signal injection attacks require repeated injections to be effective, and (2) that the regularity of normal signals admits an exceptionally accurate anomaly detection capability. Our initial experimental results tested two covert and one DOS attack that all require regular-frequency signal injections. In light of the plots of Figure 3 it is clear that any of the the three attacks used in this experiment are nearly immediately identifiable by the prolonged series of alerts, and we believe this near perfect detection capability to apply any sustained regular-frequency signal injection attack. We conclude that our approach is a promising avenue for accurate detection of an important class of CAN bus attacks.

While significant progress has been realized in our detection capabilities, tweaking and hardening the algorithm is necessary for transition to practice; e.g., understanding if/when an ECU can experience a non-malicous change of state that changes its signal frequency. Future mathematical investigations of the speed of convergence of empirical distributions will lead to theorem-driven, principled choices for currently manually-tuned parameters (e.g., thresholds, training time). Experiments with a larger number of attacks and wider variety of attacks is necessary. Finally, exploring after-market hardware/software solutions for deployment are needed for transition.

Future work will focus on extending this solution to rapidly address a wider array of vehicles. This effort will employ two different approaches: 1) improving the efficiency of "learning" the

[1]Micro averages computed by summing the number of true/false positives/negatives across the five tests, then computing rates from cumulative counts. False detection rate is defined as %alerts that are false positives.

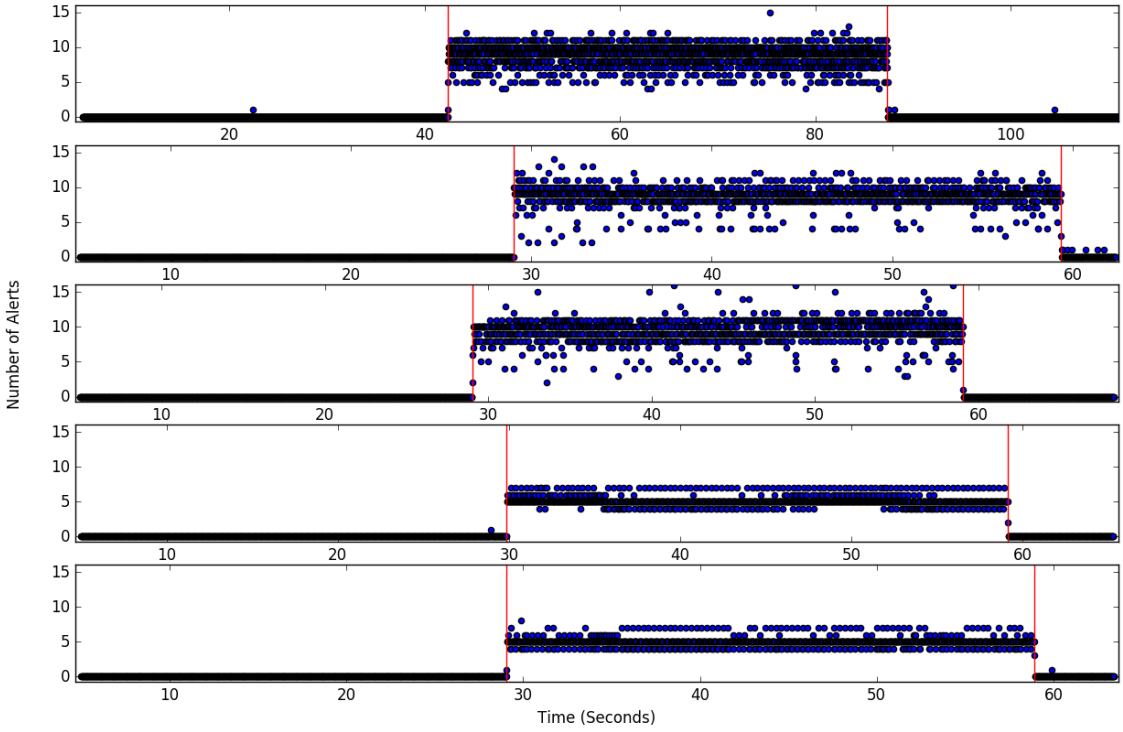

Figure 3: Results of five initial tests of the inter-signal wait-time anomaly detector depicted. (From top to bottom) first/second test runtime lights attack with engine off/on; third test is the reverse lights attack with engine off; fourth/fifth tests are the DOS attack with engine on/off. For each test, only the first five seconds of data are used to train the 120 anomaly detectors. Red vertical bars indicate the start and end of each attack. The total number of anomalies in each 0.025s interval is plotted. Micro-averaged rates are as follows: true positive rate = 0.9998, false positive rate = 0.00294, false detection rate = 0.00298.

necessary subset of CAN bus commands (PIDs) and payload values for a new vehicle (e.g., a new 2018 model) and 2) developing cyber detection algorithms that are more generalized, and thus reduce the need for exhaustive detailed knowledge of each proprietary system. The efforts to improve the learning time required for each new vehicle will focus on advances in machine learning methods coupled with the inclusion of heuristics developed from the study of a wide variety of current vehicles. One such method is a localized version of war-driving that is currently being tailored to first identify the more detrimental attack effects (e.g., engine speed) and then determine all combinations of PIDs, payloads and inter-command timings that can cause that effect.

REFERENCES

[1] Ian Broster and Alan Burns. 2003. An analysable bus-guardian for event-triggered communication. In Real-Time Systems Symposium, 2003. RTSS 2003. 24th IEEE. IEEE, 410–419.

[2] Stephen Checkoway, Damon McCoy, Brian Kantor, Danny Anderson, Hovav Shacham, Stefan Savage, Karl Koscher, Alexei Czeskis, Franziska Roesner, Tadayoshi Kohno, and others. 2011. Comprehensive Experimental Analyses of Automotive Attack Surfaces.. In USENIX Security Symposium. San Francisco.

[3] Tobias Hoppe, Stefan Kiltz, and Jana Dittmann. 2008. Security Threats to Automotive CAN Networks – Practical Examples and Selected Short-Term Countermeasures. Springer Berlin Heidelberg, Berlin, Heidelberg, 235–248. DOI: http://dx.doi.org/10.1007/978-3-540-87698-4_21

[4] Karl Koscher, Alexei Czeskis, Franziska Roesner, Shwetak Patel, Tadayoshi Kohno, Stephen Checkoway, Damon McCoy, Brian Kantor, Danny Anderson, Hovav Shacham, and others. 2010. Experimental security analysis of a modern automobile. In Security and Privacy (SP), 2010 IEEE Symposium on. IEEE, 447–462.

[5] Tsutomu Matsumoto, Masato Hata, Masato Tanabe, Katsunari Yoshioka, and Kazuomi Oishi. 2012. A method of preventing unauthorized data transmission in controller area network. In Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th. IEEE, 1–5.

[6] Michael Müter and Naim Asaj. 2011. Entropy-based anomaly detection for in-vehicle networks. In Intelligent Vehicles Symposium (IV), 2011 IEEE. IEEE, 1110–1115.

[7] Michael Müter, André Groll, and Felix C Freiling. 2010. A structured approach to anomaly detection for in-vehicle networks. In Information Assurance and Security (IAS), 2010 Sixth International Conference on. IEEE, 92–98.

[8] Hendrik Schweppe and Yves Roudier. 2012. Security and privacy for in-vehicle networks. In Vehicular Communications, Sensing, and Computing (VCSC), 2012 IEEE 1st International Workshop on. IEEE, 12–17.

[9] Ivan Studnia, Vincent Nicomette, Eric Alata, Yves Deswarte, Mohamed Kaâniche, and Youssef Laarouchi. 2013. Survey on security threats and protection mechanisms in embedded automotive networks. In Dependable Systems and Networks Workshop (DSN-W), 2013 43rd Annual IEEE/IFIP Conference on. IEEE, 1–12.

[10] Marko Wolf, André Weimerskirch, and Thomas Wollinger. 2007. State of the art: Embedding security in vehicles. EURASIP Journal on Embedded Systems 2007, 1 (2007), 074706.

View publication stats