Wireless Physical Layer Authentication via Fingerprint Embedding

Paul L. Yu, Gunjan Verma, and Brian M. Sadler

ABSTRACT

Authentication is a fundamental requirement for secure communications. In this article, we describe a general framework for fingerprint embedding at the physical layer in order to provide message authentication that is secure and bandwidth-efficient. Rather than depending on channel or device characteristics that are outside of our control, deliberate fingerprint embedding for message authentication enables control over performance trade-offs by design. Furthermore, low-power fingerprint designs enhance security by making the authentication tags less accessible to adversaries. We define metrics for communications and authentication performance, and discuss the trade-offs in system design. Results from our wireless software-defined radio experiments validate the theory and demonstrate the low complexity, practicality, and enhanced security of the approach.

FINGERPRINTING RADIO WAVEFORMS

We begin with an overview of intrinsic and intentionally embedded fingerprinting, and discuss the relationship with identification, communications, secrecy, and authentication. We then introduce a method of embedding fingerprints for wireless authentication that overcomes the deficiencies of using intrinsic fingerprints to identify the transmitter. Furthermore, the method has very small bandwidth requirements compared to traditional message authentication codes, making it a natural candidate for bandwidth-constrained environments such as mobile ad hoc networks (MANET).

INTRINSIC FINGERPRINTS

A fingerprint is literally the impression of a fingertip, but more broadly is a characteristic that identifies. This is often associated with an intrinsic property of uniqueness, or at least uniqueness viewed as a realization of a random process with structure. For example, identifying humans via biometrics now includes fingerprints, iris scans, DNA, voice features, and behavioral patterns [1]. Several investigators have considered applying these ideas to radio transmissions, including identification of radios based on signal transients [2], and study of vulnerability of these methods to impersonation [3]. In wired communications, identification of Ethernet cards has been demonstrated [4]. To be of practical use, fingerprints should be easily measurable with a sensor that is convenient and technologically feasible, and be robust to measurement noise. In addition, security features such as tamper resistance may be desirable, but these are not necessarily inherent in the fingerprint.

While these examples illustrate fingerprints derived from intrinsic structure, one may also derive a fingerprint from intrinsic randomness. For example, some wireless physical layer security techniques exploit the fact that a realization of a fading communications channel between any two agents is unique. Therefore, the channel realization may be used as a means of identifying the transmitter [5]. However, the channel properties are outside of our control, and can be noisy and rapidly time-varying, placing limits on the ability to systematically design for security.

EMBEDDED DEVICE FINGERPRINTS

Control of performance can be achieved by purposefully embedding a fingerprint in a device in a designed way so that each device can be uniquely identified. Now, in addition to the above characteristics of goodness, a good fingerprint will have strong security features, including the ability to defeat cloning and tampering. That is, a good fingerprint is not only unique and identifiable, but also hard to copy (spoof) or remove. Thus, for example, a manufacturer can label and recognize each individually manufactured device, while at the same time making it difficult to produce a counterfeit that cannot be differentiated from the genuine original. Needless to say, this has important implications for commercial enterprise.

Fingerprint embedding into devices is, of course, dependent on the specifics of the device, and so can take many forms. In an effort to make fingerprints uncloneable, some manufacturers have purposefully injected randomness into the manufacturing process to create unique characteristics [6]. Such an intrinsic fingerprint

The authors are with the U.S. Army Research Laboratory.

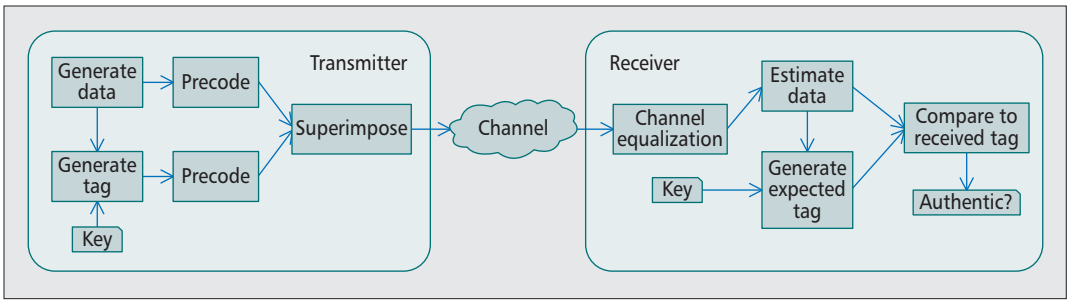

Figure 1. Physical layer authentication system diagram. The transmitter generates a data-dependent authentication tag and superimposes it with the data. The receiver estimates the data and generates the corresponding expected authentication tag. A matched filter test is performed to determine the presence of the tag in the signal and validate the transmitter's identity [8, Fig. 1].

Authentication defends against message tampering, and enables trust to be established between users. This is especially critical in wireless communications using an open and shared medium, where an adversary can eavesdrop, spoof, and jam.

can be measured and cataloged, like a serial number. The inherent randomness that is exploited for the intrinsic fingerprint may be uncontrollable at the micro-scale, and thus it may be very difficult or impossible to manufacture a similar device (a clone) with a prespecified fingerprint.

Embedded Fingerprints for PHY Authentication

Because a good fingerprint uniquely and securely identifies its source, it is an ideal candidate to convey authentication, the process of validating the identity of a message source while rejecting impersonation attacks by an adversary. Authentication defends against message tampering, and enables trust to be established between users. This is especially critical in wireless communications using an open and shared medium, where an adversary can eavesdrop, spoof, and jam.

Conventional authentication schemes transmit both the data message and a separate authentication message, referred to as a tag [7]. Each authentication tag is dependent on the associated data, and a secret key that is shared only between the transmitter and receiver. The tag generator employed at the transmitter is often a cryptographic hash function, the input of which is the key and the packet message, and its output is the tag. Such functions are highly nonlinear and difficult to invert, so an adversary cannot easily recover the key given the message and the associated tag.

In the conventional approach, the authentication tag is appended to the message, and both are transmitted at the same power. This has two disadvantages. One, the tag reduces spectral efficiency because it is time-multiplexed with the data. Two, the tag is available at high fidelity at the receiver.

Consequently, in this case, authentication security is solely predicated on encryption via the hash function, and in principle is susceptible to discovery if an adversary has sufficient computational resources. This motivates the use of an embedded fingerprint to carry the authentication tag.

An embedded fingerprint can be designed so that its bandwidth requirements are low, and its recovery is difficult for the adversary [8–10]. That is, we can arbitrarily decrease the ability of the adversary to observe the authentication tag by lowering the power of its embedding. As we show later, this leads to uncertainty about a secret key that is not readily defeated by an increase in the adversary's computational ability. Thus, embedding provides additional security and, unlike the conventional authentication approach, does not solely rely on cryptographic security. However, there is a design trade-off, because lowering the tag embedding power also weakens the ability of the intended receiver to authenticate valid packets. As we show below, good performance trade-offs are readily achieved in a software-defined radio (SDR) operating in a fading environment. The design trade-offs are fully characterized in [8–10].

Embedded Authentication Framework

The authentication system is diagrammed in Fig. 1. The transmitter (Alice) generates the authentication tag using the message and a shared secret symmetric key. The fingerprint is embedded in the transmission by adding the low-power tag to the message signal. The receiver (Bob) decodes the message and locally generates the expected authentication tag with his copy of the secret key. Finally, Bob validates the message as authentic if he detects the presence of the tag in the received signal.

Authentication System: Transmitter

For ease of presentation, we consider the case where Alice and Bob are each using single-antenna wireless devices communicating over a single carrier frequency [9]. The method readily generalizes to multi-carrier [10] and multi-antenna multiple-input multiple-output (MIMO) cases [8]. Alice has a message S to give to Bob, with whom she shares a secret key k. She first generates an authentication tag using a tag generating function, $T = g(S, k)$, which is based on a cryptographic hash function.[1] Then she transmits the weighted superposition of the message and tag signals,

$$X = \rho_S S + \rho_T T.$$

In order to make it difficult for the adversary to recover information about the authentication

[1] A cryptographic hash function is easy to compute, but infeasible to invert [7]. Furthermore, it is infeasible to find two messages that result in the same hash. RIPEMD-160 and SHA-2 are two examples.

tag, and to minimize self-interference, we set $\rho_S \gg \rho_T$ (i.e., the fingerprint has proportionally very low power). For any value of $\rho_T \in [0,1)$ we choose ρ_S so that the expected power of X remains constant, so we can regard ρ_S and ρ_T as power allocation percentages between the message and the fingerprint.

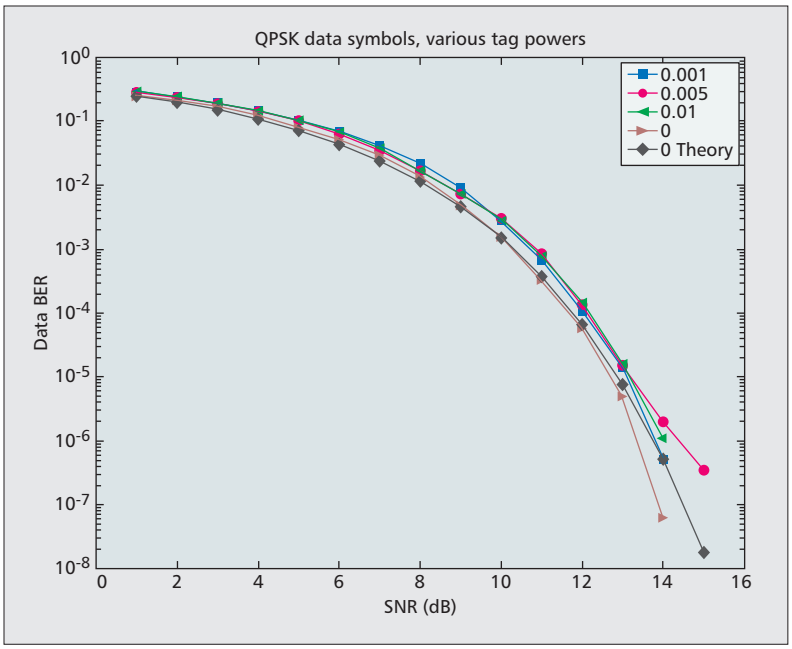

Figure 2. SDR experiment results. Low-powered authentication has minimal impact on the data bit error rate. Tag powers range from 0.1 to 1 percent of the transmit power. The 0 tag power case corresponds to the data only situation where no authentication is transmitted. The results show good agreement with the overlaid theoretical additive white Gaussian noise (AWGN) BER curve labeled 0 Theory.

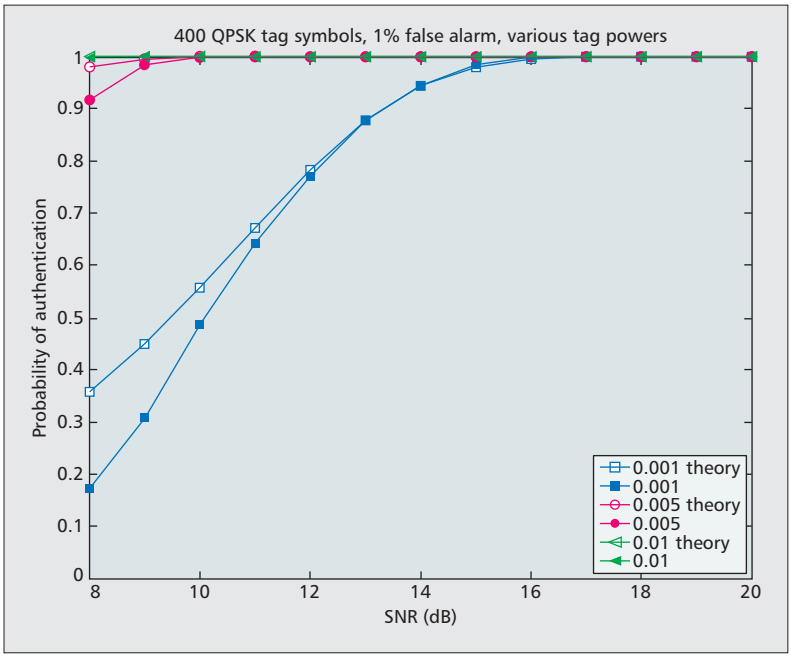

Figure 3. SDR experiment results. Authentication probability for various tag powers (from 0.1 to 1 percent of the transmit power). Packets contain 400 QPSK symbols, and the authentication false alarm probability is 1 percent. Higher-powered tags have high authentication performance.

Authentication System: Receiver

The receiver processing and authentication steps are shown in Fig. 1. Bob reverses the effect of the channel through equalization, and proceeds to demodulate and decode the message as usual. The fingerprint need not be embedded in pilot symbols so as to avoid interference with channel estimation and equalization. Because the authentication tag T is data-dependent, the receiver is unable to remove it prior to decoding the data, and thus it acts as interference during data recovery. However, by keeping the tag power relatively low, we show that message recovery is minimally impacted.

Referring again to Fig. 1, having estimated the data, the receiver can now proceed to complete the authentication process. Bob uses his shared secret key and the received data to generate the tag he expects to see. After removing the recovered data from the received signal to form a residual signal, the receiver determines the presence or absence of the authentication tag via a matched filter test comparing the expected tag and the residual signal [9, eq. 20].

The performance of the authentication test is determined by the energy of the authentication tag, which is under our design control through the tag length and the tag power allocation. Analysis and experiments show that for even moderate message packet sizes, the correlation test statistic is well approximated as having a Gaussian distribution, so the test threshold is readily set to achieve desired false alarm or detection probability [8–10].

The received tag is in the residual signal, and so is noisy. We rely on the matched filter test to overcome the noise to achieve reliable authentication. Cryptographic hash functions are used so that tags generated for distinct messages are generally far apart in Hamming distance, even when the messages are close in Hamming distance. This feature ensures strong resistance to spoofing attacks [7].

Theory and Experiments

We quantify the performance of our embedded fingerprinting approach by considering the effect of the fingerprint on the data demodulation, and the ability of the receiver to authenticate packets. In the following we present single-antenna single-carrier experiments using NI-USRP SDRs [11, 12], and compare these with theoretical predictions [9]. We use two USRP1 devices under MATLAB control [13], at a frequency of 2.39 GHz, employing quadrature phase shift keying (QPSK) modulation for the data and the tag.[2] The two radios are placed about 15 ft apart in an office building with many scatterers in the scene. By scaling the power of the transmitter, various receive signal-to-noise ratio (SNR) levels are attained, and between 25,000 and 30,000 packets are transmitted at each level under test.

[2] In this experiment we employ symbol synchronous messages and tags, which simplifies the implementation. However, the tag may be inserted into the transmit waveform in a variety of ways. For example, it need not be additive, which is a topic for further research.

Impact on Data BER

The superimposed tag takes power away from the data signal and acts as interference to data modulation. When the tag is superimposed at low power, the interference to data demodulation may be modeled as an increase in noise (i.e., as a decrease in data SNR). For example, suppose that a given channel realization yields 10 dB received SNR. If the tag uses 1 percent of the power, the data SNR becomes 9.94 dB. Hence, the data bit error rate (BER) is essentially unchanged at such a low authentication power, and the interference caused by the tag is minimal. Simulation and experimental results confirm this to be the case [8–10].

Figure 2 shows the impact of the authentication on the data BER for an over-the-air SDR experiment. For the tag power allocations tested (0.1, 0.5, and 1 percent), the BER curves are, for practical purposes, coincident. The theoretical and experimental additive white Gaussian noise BER curves (with no tag present) are also shown for comparison, validating the experimental results. We next show that while the change in data BER is essentially negligible for small tag powers, these low tag power levels are sufficient for robust authentication.

Authentication Performance

Figure 3 shows the experimental and theoretical authentication performance for various tag powers. The packets contain 400 QPSK data symbols and a corresponding 400 QPSK symbol tag, with tag power ranging from 0.1 to 1 percent of the transmit power. The authentication test threshold was set to achieve a 1 percent false alarm probability. This figure shows the effect of changing the tag power while holding the packet length constant. As previously discussed, authentication performance is improved by increasing the tag energy. Additional experiments show that modifying the packet length also yields very good agreement between theory and experiment. Other design variations include spreading the tag over multiple packets.

Security

In this section we quantify the benefit of transmitting a low-power tag in terms of how well it preserves the key's secrecy. Although Eve does not initially have the secret key k shared by Alice and Bob, she gains key information by observing their communications [14]. If she has complete knowledge of k and the tag generating function, she is able to impersonate Alice by generating legitimate tags for her messages. The protection of the key is therefore crucial to the security of the authentication system.

In the following we quantify the effort required for the adversary to learn the secret key from embedded fingerprints. We assume that Eve, just like Bob, is able to recover the data from her observations without error. Furthermore, we assume that Eve knows the tag generating function that defines the dependence of the tag on both the message and the secret key. Knowledge of the tag generation function implies knowledge of the set of possible keys, for example, the set of integers between 0 and $2^{32} - 1$, corresponding to 32-bit keys.

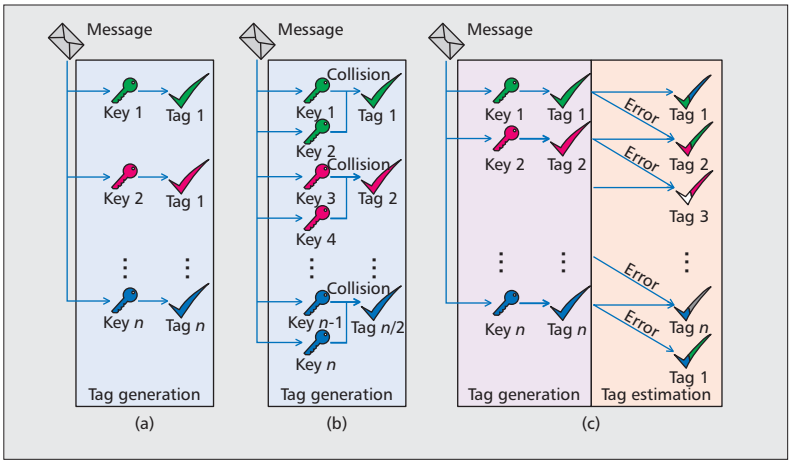

Figure 4. Illustration of key equivocation with embedded fingerprints. The tag generation is considered for a single fixed message for each panel: a) each key maps the message to a unique tag; knowledge of the message and tag leaves no uncertainty as to which key was used, and hence key equivocation is zero (key equivocation = 0 bits, noise-free tags); b) exactly two keys collide to map the message to the same tag, leading to uncertainty about which key (of the same color) was used even when the message and tag are known without error (key equivocation = 1 bit, noise-free tags; c) although the tag mapping is unique, exactly two tags can result in the same tag estimate, again leading to uncertainty about which tag was actually transmitted (key equivocation = 1 bit, noisy tags). Hence, positive key equivocation arises in both b and c.

The key equivocation, or conditional entropy [15], is a measure of Eve's uncertainty about the secret key given her observations and presumed infinite computational resources. This bounds the ability of the adversary to attack the authentication system [14]. Equivocation is non-negative and bounded from above by the entropy of the secret key. Zero key equivocation implies that Eve has no uncertainty about the key. Having no uncertainty means that there is only one key that fits the observations, although it may require a great deal of computation to ascertain its value. Zero key equivocation is the worst case for security because brute-force attacks are guaranteed to succeed (although the time required may be very long).

Figure 4 illustrates how positive key equivocation arises and how noise increases protection of the secret key. We consider the set of possible tags that can be generated from a fixed message and a set of n keys. In Fig. 4a, each key maps the message to a unique tag. Knowledge of the message and tag (as in a conventional scheme) leaves no uncertainty as to which key was used, and hence key equivocation is zero. In Fig. 4b, each key collides with exactly one other key to map the message to the same tag. Therefore even with knowledge of the message and tag, there is uncertainty about which of the colliding keys was used (e.g., key 1 vs. key 2). Finally, in Fig. 4c, we consider the effect of tag estimation on the zero key equivocation system in Fig. 4a, as occurs when the tag is embedded and only a noisy observation is available to the adversary. Estimation errors lead to uncertainty about which tag was transmitted, and hence which key was used.

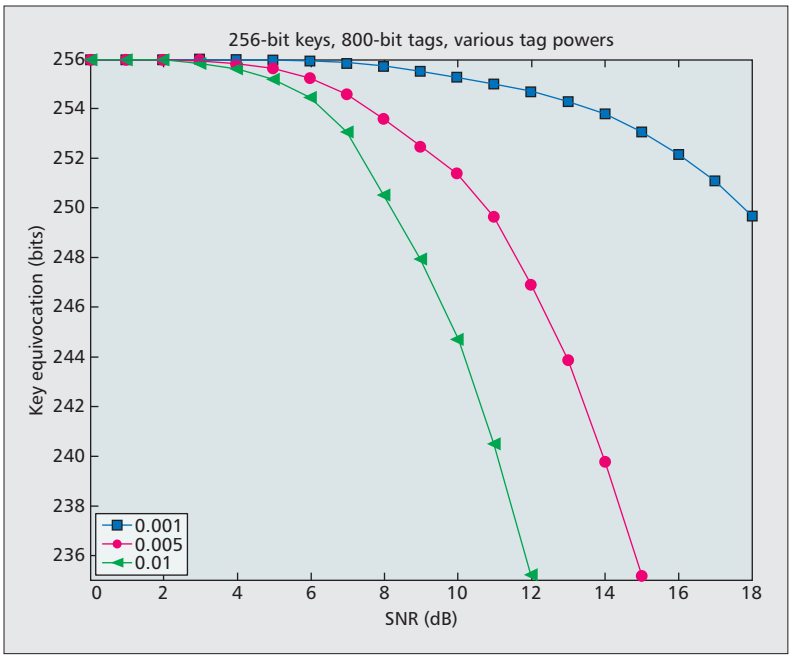

Figure 5. SDR experiment results. Key equivocation for various tag powers (from 0.1 to 1 percent of the transmit power). Low-powered tags have high key equivocation.

Intuitively, the more uncertainty the adversary has about which tag is present in the signal, the more uncertainty she has about the secret key. It then follows that lowering the tag power to present only a very noisy tag to the adversary results in high key equivocation. Although the same signal (modulo channel effects) is presented to both the intended receiver and adversary, it should be noted that the two parties have very different goals. The intended receiver has a detection problem: deciding if the tag corresponding to the received data and his key is present (corresponding to 1 bit of information). The eavesdropper has the much harder estimation problem: determining the transmitted tag and then the secret key (requiring $\log(n)$ bits of information). Because Bob is only trying to make a 1-bit decision, we can set the authentication power quite low without significantly degrading his authentication performance, while at the same time severely limiting the ability of Eve to extract key information.

Figure 5 shows experimental results for the key equivocation at various tag powers. The key equivocation is calculated based on the observed tag BER for each SNR. This figure depicts the case where key k has 256 bits and authentication tag T has 800 bits. We assume that there is zero key equivocation in the noiseless case, that is, each (message, tag) pair is associated with a unique key. In terms of equivocation this is the worst case scenario, so typical results will be better than those shown in Fig. 5.

Note that higher received SNR decreases the key equivocation. Intuitively, a cleaner observation leads to less uncertainty of the tag and hence the key. For the scenarios of interest (low tag power), the key equivocation is seen to be very high as a proportion of its 256-bit maximum.

Also note that lower tag power increases the key equivocation. As with the effect of received SNR, reducing the tag power reduces the ability of the receiver to make an accurate estimate of the tag. In this example, a large increase in key equivocation is apparent when reducing the tag power from 1 to 0.1 percent. As shown in Fig. 3, reducing the tag power does impact the ability of the intended receiver to authenticate properly. Hence, a design balance is sought to achieve the desired authentication performance while maintaining a high level of security.

Conclusions

Fingerprint embedding provides a flexible framework for message authentication, increasing security by obscuring the authentication tag to the adversary and saving transmission bandwidth. The method is readily adapted to software-defined radio, with low complexity. SDR experiments validate the theory, which enables controlled design of desired system trade-offs. There is ample room for the designer to choose the appropriate operating point that balances authentication probability and key equivocation because the impact on data BER is shown to be so slight for tag powers as high as 1 percent. The tag power and length are two parameters than can be chosen to satisfy the design requirements for arbitrary modulation schemes for the fingerprint. For example, suppose that the message is 400 symbols in length and the design requires > 99 percent authentication probability, > 250 bits of key equivocation, and $< 10^{-3}$ message BER at 10 dB SNR. Then, from Figs. 2, 3, and 5, we see that setting $\rho_T = 0.5$ percent and using 400 QPSK symbols for the tag satisfy the requirements. Alternatively, one may also decrease ρ_T and increase the length of the tag to satisfy the same requirements.

References

[1] A. K. Jain, A. A. Ross, and K. Nandakumar, *Introduction to Biometrics*, Springer, 2011.
[2] B. Danev and S. Čapkun, "Transient-Based Identification of Wireless Sensor Nodes," *Proc. ACM/IEEE IPSN*, 2009.
[3] B. Danev *et al.*, "Attacks on Physical-Layer Identification," *Proc. 3th ACM Conf. Wireless Network Security*, 2010, pp. 89–98.
[4] R. Gerdes *et al.*, "Physical-Layer Identification of Wired Ethernet Devices," *IEEE Trans. Info. Forensics Security*, vol. 7, no. 4, Aug. 2012, pp. 1339–53.
[5] L. Xiao *et al.*, "Using the Physical Layer for Wireless Authentication in Time-Variant Channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, July 2008, pp. 2571–79.
[6] U. Ruhrmair, S. Devadas, and F. Koushanfar, "Security Based on Physical Unclonability and Disorder," *Introduction to Hardware Security and Trust*, M. Tehranipoor and C. Wang, Eds. Springer, 2011.
[7] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 2001.
[8] P. Yu and B. Sadler, "MIMO Authentication via Deliberate Fingerprinting at the Physical *Layer," IEEE Trans. Info. Forensics Security*, vol. 6, no. 3, 2011, pp. 606–15.
[9] P. Yu, J. Baras, and B. Sadler, "Physical-Layer Authentication," *IEEE Trans. Info. Forensics Security*, vol. 3, no. 1, Mar. 2008, pp. 38–51.
[10] —, "Multicarrier Authentication at the Physical Layer," *Proc. Int'l. Symp. on a World of Wireless, Mobile and Multimedia Networks*, 2008, pp. 1–6.
[11] Ettus Research — Product Detail, Mar. 2013; https://www.ettus.com/product/details/USRP-PKG
[12] G. Verma, P. Yu, and B. Sadler, "Physical Layer Authentication via Fingerprint Embedding Using Software-Defined Radios," *IEEE Access*, vol. 3, 2015, pp. 81–88.

[13] G. Verma and P. L. Yu, "A MATLAB Library for Rapid Prototyping of Wireless Communications Algorithms with the USRP Radio Family," U.S. Army Research Lab., tech. rep., 2013.
[14] U. Maurer, "Authentication Theory and Hypothesis Testing," *IEEE Trans. Info. Theory*, vol. 46, no. 4, July 2000, pp. 1350–56.
[15] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley-Interscience, 1991.

Biographies

Paul L. Yu [M] received B.S. degrees in mathematics and computer engineering, and a Ph.D. degree in electrical engineering. He is with the U.S. Army Research Laboratory (ARL), where his work is in the areas of signal processing, wireless communications, and network science.

Gunjan Verma received B.S. degrees in mathematics and computer science, and a B.A. degree in economics from Rutgers University, and an M.S. degree in applied mathematics from Johns Hopkins University. He is with the U.S. ARL, where his work spans information flow in complex networks, signal processing, and validation and prototyping of wireless communications algorithms using software-defined radios.

Brian M. Sadler [S'81, M'81, SM'02, F'07] received B.S. and M.S. degrees from the University of Maryland, College Park, and his Ph.D. degree from the University of Virginia, Charlottesville, all in electrical engineering. He is a Fellow of the ARL in Adelphi, Maryland. He is an Associate Editor for *EURASIP Signal Processing*, was an Associate Editor for *IEEE Transactions on Signal Processing* and *IEEE Signal Processing Letters*, and has been a Guest Editor for several journals including *IEEE JSTSP*, *IEEE JSAC*, *IEEE Signal Processing Magazine*, and the *International Journal of Robotics Research*. He is a member of the IEEE Signal Processing Society Sensor Array and Multi-Channel Technical Committee, and Co-Chair of the IEEE Robotics and Automation Society Technical Committee on Networked Robotics. He received Best Paper Awards from the Signal Processing Society in 2006 and 2010. He has received several ARL and Army R&D awards, as well as a 2008 Outstanding Invention of the Year Award from the University of Maryland. His research interests include information science, networked and autonomous systems, sensing, and mixed-signal integrated circuit architectures.