TRIFECTA: Security, Energy Efficiency, and Communication Capacity Comparison for Wireless IoT Devices

Shreyas Sen
Purdue University

Jinkyu Koo
Purdue University

Saurabh Bagchi
Purdue University

The widespread proliferation of sensor nodes in the era of the Internet of Things (IoT), coupled with increasing sensor fidelity and data acquisition modality, is expected to have generated 3+ Exabytes of data per day by 2018. Since most of these IoT devices will be wirelessly connected at the last few feet, wireless communication is an integral part of the future IoT scenario. The ever-shrinking size of unit computation (Moore's law) and continued improvements in efficient communication (Shannon's law) are expected to harness the true potential of the IoT revolution and produce a dramatic societal impact. However, reducing the size of IoT nodes and the lack of significant improvement in energy storage density leads to reduced energy availability. Moreover, smaller size and less energy means fewer resources available for securing IoT nodes, making the energy-sparse low-cost leaf nodes of the network prime targets for attackers. In this article, we survey six prominent wireless technologies with respect to three dimensions: security, energy efficiency, and communication capacity. We point out the state of the art, open issues, and the road ahead for promising research directions.

We see the following five trends in connected computing.

First, cheap ubiquitous computing is leading to *smart things*. Through five decades of continued scaling, following Moore's law, the size of unit computing has gone to virtually zero. Starting with mainframe computers in the '60s, which used to be the size of a room, we've seen continuous reduction in the size of a computer. We've seen computers progress through the mini, the workstation, and the PC down to laptops in the 2000s (see Figure 1). The 2010s have been dominated by mobile devices (e.g., smartphones). By the year 2020, the size of unit (meaningful) computation will be so small that it will be barely visible. This will enable cheap, ubiquitous computation all around us, incorporated into everyday things like wearables, household devices, and mobile payment devices. The ability to incorporate significant computation in an almost invisible footprint is transforming everyday objects into smart things.

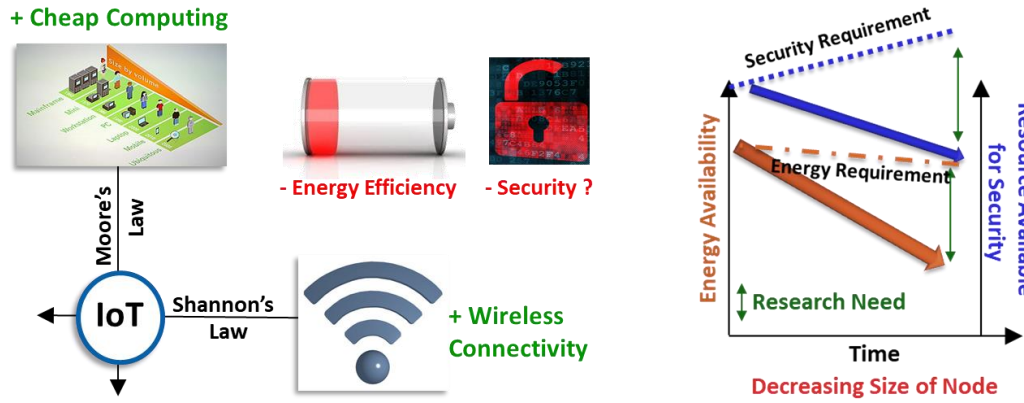

Figure 1. The Internet of Things (IoT) as the junction of Moore's law and Shannon's law. A big bottleneck for ubiquitous computing using IoT nodes is reduced energy availability and increased security vulnerabilities in leaf nodes as the size of the node decreases.

Second, cheap wireless connectivity is leading to connected things. The emergence of the Internet as a household commodity worldwide, coupled with tremendous progress in commoditization of wireless connectivity (especially cellular 5G and wireless LANs), means that billions of things can now be wirelessly connected to the Internet.

Third, smart connected things are leading to the Internet of Things (IoT). The emergence of cheap computing following Moore's law is enabling smart things, and the emergence of cheap wireless connectivity following Shannon's law (see Figure 1) is creating smart connected things. At present, we are standing at the crossroads of smart and connected IoT, which is quickly transforming human lives. The number of Internet-connected devices already passed the number of human beings on the planet in 2009 and is increasing exponentially. Cisco estimates that by 2020, there will be 3.4 devices connected to the Internet per person.

Fourth, smaller size and similar energy density are leading to lower available energy. Though the size of unit computation is falling fast, the energy storage or battery technology is improving only very slowly, leading to a reduced amount of available energy in smaller nodes. Due to its small footprint, the size of the battery included in such sensor nodes is limited. Moreover, including a battery means increased deployment cost and, more important, maintenance cost (to change the battery periodically). Since the electronics' lifetime is generally significantly higher than battery lifetime, it is desirable to develop net-zero-energy sensor nodes, which perpetually run on harvested energy. However, the trend is to pack more and more functionality even into energy-constrained nodes, and they need to communicate wirelessly. This leads to an energy gap and calls for significant improvement in energy efficiency for computing and communication in energy-constrained nodes. Some recent software frameworks for healthcare applications are highlighting the software blocks responsible for greatest energy drain[16] and providing software abstractions for more energy-efficient application development[15].

Finally, smaller size and lower energy are leading to lower resource availability for security. It is well known that the security of a network is often only as good as its weakest link. Energy-

sparse, size-constrained IoT nodes have limited resources to guarantee strong security and hence often are the weakest link in the end-to-end system. While the resources available for security are decreasing (see the bottom of Figure 1), with reduced size, the security requirements of these leaf nodes are increasing, creating a strong need for research in lightweight, resource-constrained security technologies.

In many of the compelling application areas, the security of the communication channel is of primary importance, including the possibility of eavesdropping (i.e., loss of confidentiality) and denial of service (i.e., loss of availability).[1] The two concerns that have traditionally been looked at for this class of systems are energy efficiency and communication capacity.[2] In this article, we analyze prominent wireless technologies for the IoT with respect to the three dimensions: security, energy efficiency, and communication capacity. These dimensions are of course interrelated; e.g., an otherwise energy-efficient system may become unusable if it needs cryptographic protocols, which are expensive on such systems.

PROMINENT TECHNOLOGIES

Multiple classes of wireless technologies—namely, wireless local area networks (WLANs: Wi-Fi and Bluetooth), sensor networks (ZigBee), near-field technologies (near-field communication and emerging high-speed proximity communication[3]), and wide-area wireless communication (LoRa)—will be compared across security, energy efficiency, and communication capacity.

Energy Efficiency and Communication Capacity

The State of the Art

Table 1 summarizes several wireless techniques in terms of the communication distance supported, typical energy efficiency (in Joules per bit), data rate (communication capacity in bits per second), and security. Figure 2 plots these technologies for energy efficiency (*y*-axis) versus the data rate (*x*-axis). It is to be noted that the maximum data rate is often limited by the US Federal Communications Commission and the standard. The communication energy efficiency varies from several from pJ/b to μJ/b—i.e., six orders of magnitude, depending on the PHY and the distance supported. A significant amount of this energy is wasted due to inflexible, worst-case radio designs.[4,17,18]

//Author: In Table 1, did the L, M, and H in the Security row stand for Low, Medium, and High, as I assumed? Also, some of the cell entries in that row have a superscript 1. What does it stand for? Is it in reference to a footnote? If so, what should that footnote be?//

Saurabh: Addressed.

Table 1. Comparison of state-of-the-art wireless techniques for Internet of Things nodes.

	Proximity	NFC	ZigBee	BT	WiFi	LoRa
Distance	1 mm	10 cm	10–100 m	10–100 m	30–50 m	~km
Data rate	8–32 Gbps	0.1–0.8 Mbps	0.02–0.2 Mbps	0.8–2.1 Mbps	300 Mbps (11g) 7 Gbps (11ac, 11d)	200 Kbps
Energy-efficiency	4 pJ/b	1-50 nJ/b	5 nJ/b	15 nJ/b	5 nJ/b	1 μJ/b
Security[1]	High[3]	Medium	Low[4]	Low	Medium or high[2]	Relatively unknown

1. The security level is given as “H” (high), “M” (medium), and “L” (low). This represents not the theoretically achievable security, but rather the security of most common deployments of the technology.

2. With state-of-the-art software extensions to the basic standard, the level “H” is achieved.

3. The “H” level is due in large part to the physical proximity of the communicating devices.

4. See for example: “ZigBee Exploited the Good, the Bad, and the Ugly,” Tobias Zillner (Cognosec GMBH), BlackHat 2015.

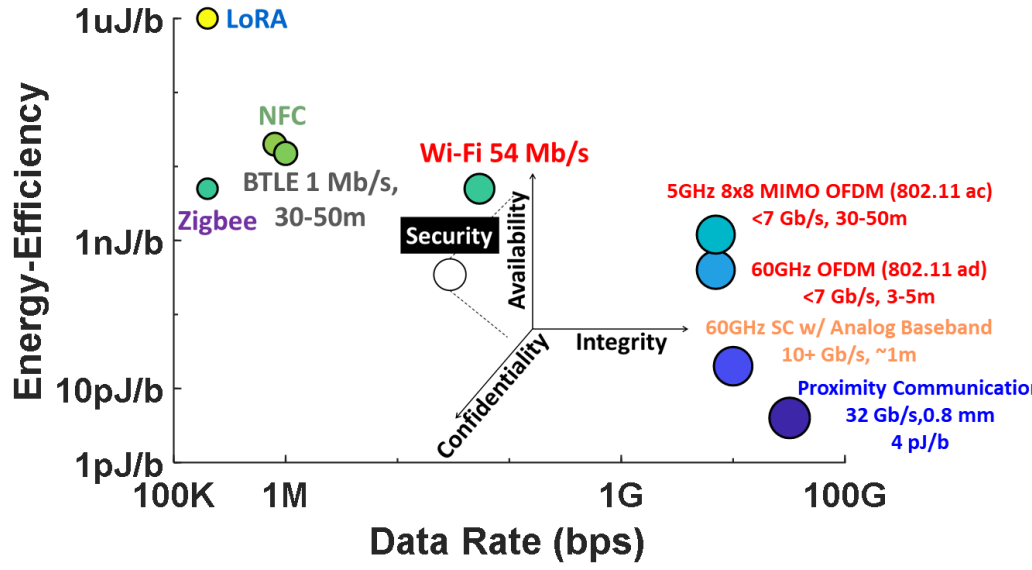

Figure 2. State-of-the-art energy efficiency versus the communication data rate of prominent wireless IoT PHYs. The size of the circle represents the strength of security, which consists of confidentiality, availability, and integrity. The gap on the bottom left motivates the research need for low-speed, reliable, yet highly efficient secure communication techniques.

Energy Gap

Current battery technology supports enough energy for low-performance communication, and hence we are seeing a plethora of commercial low-performance battery-operated IoT devices. However, mobile devices are severely energy constrained for both battery-operated high-performance devices and energy-harvesting low-performance devices. A typical smartphone battery holds 5 to 10 watt-hours of energy. Communicating 10 Gbps of data (e.g., 4K video, 30 fps, RGB, and 12-bit color depth = 9.56 Gbps raw) at 1 nJ/b leads to 10 watts of power. Hence, the mobile battery runs out within an hour, barely supporting such communication, let alone processing and display. Similarly, for energy-harvesting devices, solar harvesting lends tens of mWs of power in favorable outdoor conditions. However, for all other modalities (e.g., indoor-lighting, vibration, thermal, and radio frequency harvesting), typical harvested power falls in the range of 50 to 200 µW. A sensor node trying to communicate 1 Mbps (e.g., compressed or intermittent video) with an energy efficiency of 1 nJ/b will consume 1 mW just for the communication portion. This highlights the energy gap present for current IoT sensor nodes. An order-of-magnitude improvement in communication energy efficiency will open up many applications of ubiquitous connected IoT nodes.

Security Considerations

As we gradually move toward using some of the smart devices for critical operations, security will become a primary driver for which devices win out in the marketplace. We are already seeing some such uses around us, such as in mobile payment systems (Google Wallet, Apple Pay,

and Samsung Pay) and wearable healthcare devices that monitor for critical health signals (such as heart rate and VO2 level) and, in the case of critical indicators, communicate to a health provider. We survey here some of the successes and challenges for securing the wireless technologies under discussion here. We also discuss some of the unique aspects of security in this domain. In all of this, it is important to keep in mind that security should be considered as improving the state of affairs on one or more of three axes (see Figure 2): confidentiality (of the information being stored or exchanged), integrity (of the data being stored), and availability (being able to access the device and its stored state). Also, the security achieved is hardly ever zero or one for any of these axes, but rather is on a sliding scale.

Geographical Proximity as an Aid

Technologies that operate in very close proximity, such as NFC with < 20 cm range, rule out most man-in-the-middle (MITM) attacks. A typical MITM attack scenario is as follows, where Mallory, an attacker, is interposing herself between the communication of two legitimate parties, Alice and Bob. This kind of MITM attack, unique to our proximal-wireless-communication scenarios, is possible even when cryptography is being used, due to the ability of the attacker to intercept the communication.

1. Alice sends her public key to Bob, but Mallory can intercept it. Mallory sends Bob her own public key for which she has the matching private key. Now Bob wrongly thinks that he has Alice's public key.
2. Bob sends his public key to Alice, but Mallory can intercept it. Mallory sends Alice her own public key for which she has the matching private key. Now Alice wrongly thinks that she has Bob's public key.
3. Alice sends Bob a message encrypted with Mallory's public key, but Mallory can intercept it. Mallory decrypts the message with her private key, keeps a copy of the message, re-encrypts the message with Bob's public key, and sends the message to Bob. Now Bob wrongly thinks that the message came directly from Alice and has no indication that the message has been intercepted and decrypted.
4. Similar to step 3 above, when Bob sends Alice a message, Mallory can again decrypt it and optionally modify it, before passing it on to Alice, pretending that it came from Bob.

This kind of MITM attack can be mitigated if Alice and Bob have a visual connection due to geographical proximity and can prove to each other's devices that there is such proximity. This typically requires entering a secondary authentication token that appears on both devices, such as a long random PIN. It is in the choice of the secondary authentication mechanism that the capability of the device will become a crucial factor. For example, if the device has an output display, then the PIN can be displayed; if the device has a touch sensor, then Alice and Bob can be asked to authenticate by physically touching the other person's device.

Isolation and Abstraction

Isolation of different hardware and software modules has been considered a key building block for secure systems in the traditional desktop and server world. This means that there are boundaries to what each hardware or software module can access (e.g., only some parts of the device's memory), and thus, if one module is compromised, the entire system does not get compromised. In the domain under consideration here, such isolation may or may not be possible depending on the specific device's price point. One example where such isolation is widely used today is in smartphones. In most smartphones, there is the relatively powerful main processor and a separate baseband processor.[5] The baseband processor runs the radio control functions, which have real-time requirements, and therefore a real-time OS runs on the baseband processor. However, due to the proprietary nature of the software stack on it, there are often security vulnerabilities found in it.[6] The software on the main processor trusts the software running on the baseband processor, and thus the vulnerability can spread. Thus, we see that despite isolation, if the separation is not enforced, security breaches occur. Therefore, the correct design point is that whenever isolation is possible, either in hardware or software, then enforcement of the separation boundary is needed. Some recent efforts with low-end embedded devices[7,8] are showing how it is possible to en-

force isolation with limited hardware support, mainly through software techniques. We acknowledge, however, that for many low-end smart things, such isolation will be infeasible, and therefore systems must be built with an acknowledgment of the vulnerable nodes and understanding of their spread potential. One possible mitigation action that has found success in healthcare systems is the use of federated identity management whereby resources are managed by multiple organizations and whenever there is a transition from one to another federated authentication and authorization protocols are invoked[14].

Out-of-Band Mechanism for Security

An interesting interplay between multiple technologies happens in this space to provide increased security. Many security protocols rely on some out-of-band (OOB) mechanism for exchanging some critical information, which helps secure a communication channel. In authentication, OOB refers to utilizing two separate networks or channels, one of which is different from the primary network or channel, to simultaneously communicate between two parties or devices for identifying a user. For example, a cellular network is commonly used for OOB authentication. An example of OOB authentication is when an online-banking user is accessing her online bank account with a login and a one-time password (OTP) is sent to her mobile phone via SMS (short message service). The primary channel would be the online login screen where the user enters her login information and the OTP sent through the OOB channel.

In our domain, oftentimes there is a clear OOB situation, which is humans interacting through their respective devices.[9] This naturally allows certain levels of trust to be established among the communicating individuals. With the right security protocol, this trust can be transferred to devices that belong to the users, enabling two devices to establish a trusted communication channel that reflects the existing trust their users place on one another. A typical example is the pairing of two Bluetooth devices with active participation of the users. In Bluetooth's Secure Simple Pairing mode, it uses NFC for achieving security. One issue to keep in mind here is that the devices should be reasonably time synchronized, say to within tens of milliseconds. Much more accurate time synchronization has been demonstrated even for ad hoc wireless networks.[1]

THE ROAD AHEAD

Communication Capacity and Energy Efficiency

In this data-driven IoT revolution, workloads, operating conditions, and computation and communication demands on distributed and connected devices will undergo large dynamic ranges of several orders of magnitude. Energy-constrained IoT nodes will demand the highest-possible energy efficiency across the entire range of operation under changing contexts. A context could be defined as channel conditions, applications, latency, quality of service, data rate requirements, battery condition, or process variation, among others. Current systems are typically overdesigned to handle all possible contexts, which creates an unfavorable tradeoff between fidelity and power efficiency. Learning from nature, we notice that a human brain continuously adapts to its surroundings to perform more efficiently. It also self-learns[10] the optimum ways with experience. Similarly, in context-aware communication, a smart IoT device understands its own context and adapts itself on the fly for optimal energy efficiency and performance. Such context-aware communication could be divided into two distinct categories—namely, intra-PHY[11] and inter-PHY adaptation—as described in "Context-Aware Energy-Efficient Communication for IoT Sensor Nodes."[2] In brief, the former means adapting within one physical-layer wireless communication channel, while the latter involves multiple such channels.

Along with context awareness, innovative technologies specific to emerging applications can enable orders-of-magnitude improvement in both communication capacity and energy efficiency, even simultaneously. As an example, recently developed capacitive-proximity communication[3] provides a wire-like data rate (32 Gbps) and energy efficiency (4 pJ/b) without a physical wire, and enables a >100× benefit over short-range millimeter-wave communication, allowing high-speed transfer (e.g., fast video and photo downloading from smartphone to laptop just by placing it on top of the laptop, without connecting a wire). Another example is human-body communica-

tion,[12] which utilizes the conductive properties of the human body to connect wearables and implantables, reducing body-area-network connectivity energy by >100× while improving privacy, as the signals are contained mostly within the body and cannot be snooped from far away by an adversary. Similar application-specific technology developments will be needed to unlock energy efficiency of >10×.

Most important, since communication energy is often the bottleneck, it's wise to communicate "information" rather than "raw data" to and from the sensor nodes. This makes sense only if the energy cost of in-sensor information extraction (i.e., in-sensor analytics) is significantly lower than the communication energy cost and a context-dependent optimum exists between in-sensor analytics and communication. It has been shown recently[13] that by tracking this optimum energy point, a IoT wireless video sensor node can achieve a 4.3× improvement in energy efficiency.

Security

We would like to see active development of usable security solutions for this space. These security solutions will span the range in the following dimensions:

- resource consumption (computation and network communication);
- level of security (For example, does it provide protection against replay attacks? How much of a brute-force attack can it tolerate?);
- level of user intervention required (Does the user need to type in a six-digit PIN, or is only a directional pointing of the device enough?); and
- use of a trusted third party (Does the protocol require intermediation of a trusted third party?).

This is an active consideration in mobile payment systems where different product offerings keep a lot, a little, or no trusted information with the vendors, such as Google, Apple, or Samsung.

An important unmet need for security solutions is context awareness. One would want not to have to spend precious energy resources on a security protocol (which can often involve expensive network communication) if the environment is benign. For example, if there are several interfering sources of wireless communication, with potentially malicious intent, then a higher level of security posture may be warranted than in a benign environment. Two important questions are, to what extent should the system automatically infer the context, and to what extent should this be input by the user? We should take care that the cure should not become more damaging than the malaise; i.e., inferring context should not become more resource consuming than in the baseline mode.

ACKNOWLEDGMENTS

Shreyas Sen would like to acknowledge the support of Semiconductor Research Corporation (SRC Grant No. 2720.001) and US National Science Foundation (NSF CRII Award, CNS Grant No. 1657455). Saurabh Bagchi and Jinkyu Koo acknowledge the support of the US National Science Foundation through the NeTS program (grants CNS-1409506 and CNS-1409589), as well from AT&T through their Virtual University Research Initiative. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

REFERENCES

1. J. Koo, R. K. Panta, S. Bagchi, and L. Montestruque, "A tale of two synchronizing clocks," in Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, 2009, pp. 239–252.

2. S. Sen, “Invited: Context-aware energy-efficient communication for IoT sensor nodes,” in Proceedings of the 53rd Annual Design Automation Conference, 2016, p. 67.
3. C. Thakkar, S. Sen, J. Jaussi, and B. Casper, “A 32 Gb/s Bidirectional 4-channel 4 pJ/b Capacitively Coupled Link in 14 nm CMOS for Proximity Communication,” IEEE J. Solid-State Circuits, vol. PP, no. 99, pp. 1–15, 2016.
4. S. Sen, V. Natarajan, R. Senguttuvan, and A. Chatterjee, “Pro-VIZOR: process tunable virtually zero margin low power adaptive RF for wireless systems,” in Proceedings of the 45th Design Automation Conference, DAC 2008, Anaheim, CA, USA, June 8-13, 2008, 2008, pp. 492–497.
5. “Baseband Zero Day Exposes Millions of Mobile Phones to Attack,” Threatpost | The first stop for security news. [Online]. Available: https://threatpost.com/baseband-zero-day-exposes-millions-of-mobile-phones-to-attack/124833/. [Accessed: 20-Sep-2017].
6. C. Mulliner and C. Miller, “Fuzzing the Phone in your Phone,” Black Hat USA 2009.
7. A. A. Clements, N. S. Almakhdhub, K. S. Saab, P. Srivastava, J. Koo, S. Bagchi, and M. Payer, “Protecting Bare-metal Embedded Systems With Privilege Overlays,” in IEEE Symp. on Security and Privacy, pp. 289-303. IEEE, 2017.
8. “Mbed uVisor | Mbed,” Arm Mbed. [Online]. Available: www.mbed.com/en.
9. A. Kobsa, R. Sonawalla, G. Tsudik, E. Uzun, and Y. Wang, “Serial hook-ups: a comparative usability study of secure device pairing methods,” in Proceedings of the 5th Symposium on Usable Privacy and Security, 2009, p. 10.
10. D. Banerjee, B. Muldrey, X. Wang, S. Sen, and A. Chatterjee, “Self-Learning RF Receiver Systems: Process Aware Real-Time Adaptation to Channel Conditions for Low Power Operation,” IEEE Trans. Circuits Syst. Regul. Pap., vol. PP, no. 99, pp. 1–13, 2016.
11. S. Sen, V. Natarajan, S. K. Devarakond, and A. Chatterjee, “Process-Variation Tolerant Channel-Adaptive Virtually Zero-Margin Low-Power Wireless Receiver Systems,” IEEE Trans CAD Integr. Circuits Syst., vol. 33, no. 12, pp. 1764–1777, 2014.
12. S. Sen, “SocialHBC: Social Networking and Secure Authentication using Interference-Robust Human Body Communication,” in IEEE International Symposium on Low Power Electronics and Design 2016.
13. N. Cao, S. B. Nasir, S. Sen, and A. Raychowdhury, “Self-Optimizing IoT Wireless Video Sensor Node With In-Situ Data Analytics and Context-Driven Energy-Aware Real-Time Adaptation,” IEEE Trans. Circuits Syst., Volume: 64, Issue: 9, Sept. 2017, pp. 2470-2480.
14. S. Chaterji, J. Koo, N. Li, F. Meyer, A. Y. Grama, and S. Bagchi, “Federation in Genomics Pipelines: Techniques and Challenges,” Oxford Briefings in Bioinformatics, bbx102, pp. 1-10, August 2017.
15. K. Mahadik, C. Wright, J. Zhang, M. Kulkarni, S. Bagchi, and S. Chaterji. "SARVAVID: A Domain Specific Language for Developing Scalable Computational Genomics Applications." In Proceedings of the International Conference on Supercomputing (ICS), pp. 1-12, June 2016.
16. X. Liu, T. Chen, F. Qian, Z. Guo, F. Lin, X. Wang, and K. Chen. "Characterizing smartwatch usage in the wild." In Proc. of the 15th Annual International Conference on Mobile Systems, Applications, and Services (Mobisys), pp. 385-398. ACM, 2017.
17. S. Sen, D. Banerjee, M. Verhelst, and A. Chatterjee, A., "A power-scalable channel-adaptive wireless receiver based on built-in orthogonally tunable LNA." IEEE Transactions on Circuits and Systems I: Regular Papers 59, no. 5 (2012): pp. 946-957.
18. S. Sen, R. Senguttuvan, and A. Chatterjee. "Environment-adaptive concurrent companding and bias control for efficient power-amplifier operation." IEEE Transactions on Circuits and Systems I: Regular Papers 58, no. 3 (2011): 607-618.

ABOUT THE AUTHORS

//Author: Are the following bios accurate?//

Shreyas Sen is an Assistant Professor at Purdue University’s School of Electrical and Computer Engineering. Dr. Sen received his BE from Jadavpur University followed by Ph.D. from ECE,

Georgia Tech in 2011 and has over 5 years of industry research experience in Intel Labs, Qualcomm and Rambus. His current research interests include circuits/systems for IoT, Biomedical and Security. Dr. Sen is a recipient of the NSF CRII Award, AFOSR Young Investigator Award, Google Faculty Research Award, Intel Quality Award for industrywide impact on USB-C type and multiple best-paper awards. He has co-authored 2 book chapters, over 100 journal and conference papers and has 13 patents granted/pending. Dr. Sen is a Senior Member of IEEE. Contact him at shreyas@purdue.edu.

Jinkyu Koo is a research scientist at Purdue University's School of Electrical and Computer Engineering. Contact him at kooj@purdue.edu.

Saurabh Bagchi is a Professor in the School of Electrical and Computer Engineering and the Department of Computer Science (by courtesy) at Purdue University in West Lafayette, Indiana. His research interest is in distributed systems and dependable computing. He is the founding Director of a university-wide resiliency center at Purdue called CRISP (2017-present). He is the recipient of an IBM Faculty Award (2014), a Google Faculty Award (2015), and the AT&T Labs VURI Award (2016). He was elected to the IEEE Computer Society Board of Governors for the 2017-19 term. Contact him at sbagchi@purdue.edu.