

Decoding the Collisions in RFID Systems

Lei Kang*, Kaishun Wu*[†], Jin Zhang* and Haoyu Tan*

*Department of Computer Science and Engineering
Hong Kong University of Science and Technology

[†]School of Physics and Engineering
Sun Yat-sen University

{kang,kwinson,zjzj, hytan}@cse.ust.hk

Abstract—RFID has been gaining popularity due to its variety of applications, such as inventory control and localization. One important issue in RFID system is tag identification. In RFID systems, the tag randomly selects a slot to send a Random Number (RN16) packet to contend for identification. Collision happens when multiple tags select the same slot, which makes the RN packet undecodable and thus reduces the channel utilization. In this paper, we redesign the RN pattern to make the collided RNs decodable. By leveraging the collision slots, the system performance can be dramatically enhanced. This novel scheme is called DDC, which is able to *directly decode the collisions* without exact knowledge of collided RNs. In the DDC scheme, we modify the RN generator in RFID tag and add a collision decoding scheme for RFID reader. We implement DDC in GNU Radio and USRP2 based testbed to verify its feasibility. Both theoretical analysis and testbed experiment show that DDC achieves 40% tag read rate gain compared with traditional RFID protocol.

I. INTRODUCTION

Radio Frequency Identification (RFID) is an emerging wireless technology that allows tiny computer chips to be remotely powered and operated for identifiers and other information. Many applications that make use of these capabilities have been proposed, e.g., supply chain monitoring where products are labeled with tags and scanned as they are moved. RFID tags form the infrastructure of “internet of things” which enables each item in daily life to be located and managed. One widely used RFID standard is EPC Class-1 Generation-2 standard [9], which defines readers and passive tags that operate at UHF frequencies.

According to the EPC standard, multiple tags access the channel under Framed ALOHA protocol. The reader broadcasts commands to divide time into slots. Each tag randomly selects one slot to contend for identification by a packet called RN, which contains a header and a generated random number. If multiple tags contend for the channel in the same slot, a collision slot happens, and if one slot is not selected by any tag, an empty slot happens. Both collision slot and empty slot are waste. Only when one tag picks a slot, this slot is useful. The efficiency of Framed ALOHA is decided by the frame length, a larger frame length brings more empty slots and a smaller frame length causes more collisions, both of which are inefficient, only when the frame length is equal to the number of tags in the field [15], the system performance is maximized.

Previous work [2], [4], [15], [16] mainly focused on developing efficient algorithm to find the optimal frame length

or adjust the system parameters to enhance the RFID system efficiency. However, Framed ALOHA has a theoretical throughput limitation by the fact that only 36.8% slots are useful at most [10]. The efficiency of most current schemes are limited by the theoretical upper bound of ALOHA-based system’s throughput.

In this paper, we introduce DDC (directly decoding the collisions), which designs new RN pattern and is able to decode the collided RNs, so that we can breakthrough the 36.8% upper bound of ALOHA. The principle of DDC is reducing the information carried by fixed length RN to enable concurrent transmission. DDC is a new RFID reader and tag design which is modulation-independent. DDC doesn’t make any modification to the RFID MAC and introduce no overhead in the case of identifying traditional tags. If several RNs are collided in one time slot, DDC achieves the same performance as if there is only one RN is received.

DDC has the following key features.

- *It decodes collision directly*: It is different from the physical layer network coding [7], [8], where one of the collided packets must be previously known or two same packets must be collided twice. DDC requires no exact information on collided RNs and decodes the collided RNs directly by pre-defined data or signal pattern.
- *It is modulation-independent*: DDC can work in any current modulations by simply defining the data patterns, as long as it can produce periodic signals by the cycle of one or more bits. Most of modulations can achieve this, e.g. ASK, PSK and FSK.
- *It is double-compatible*: DDC makes no modification on the RFID MAC. Traditional reader can identify the new tags with no overhead and the DDC can identify the traditional tag normally.

This paper is the first to present a practical design that exploits collision decoding to increase the throughput in RFID systems. Our contributions can be summarized as follows:

- We present a novel design and algorithm for channel access contention in RFID systems, where it decodes the collisions directly by pre-defined data pattern.
- We implement our approach in GNU Radio and USRP2-based testbed to verify the feasibility of DDC.
- We evaluate our implementation in a testbed of 9 USRP2 nodes (1 reader and 8 tags). Theoretical and experimental

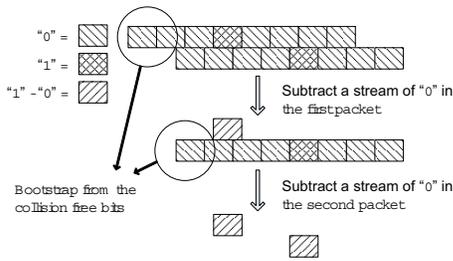


Fig. 1. Decoding Two Collided RNs.

results show that our technique improve the tag read rate by 40% compared with the EPC protocol.

This paper is organized as follows. We present how DDC works and its benefits in RFID systems in Section II. The detail of the components in DDC is presented in Section III. The implementation and experimental evaluation are illustrated in Section VI and V, respectively. The related work is summarized in Section VI. In Section VII, we conclude this paper and discuss the future work.

II. OVERVIEW OF DDC

In this section, we present an illustrative example of DDC, and how DDC outperforms the traditional EPC C1G2 UHF RFID protocol. The influence of random number pattern modification is also discussed. Finally, we will give the capacity analysis of our scheme.

A. How DDC Works

To see how DDC works, consider the example of Fig.1, where two tags pick the same time slot and transmit new designed RNs simultaneously to the reader, a collision happens. For the sake of simplicity, we assume that “1” bit and “0” bit present two distinct waveforms, e.g., sine and cosine, and all “0” bits present the same waveform pattern. We will illustrate later that nearly all modulations (ASK, PSK and FSK) can meet this requirement by simply formulating the transmitted data. Our data formulation here is to set only one “1” bit in each RN. It is worth pointing out that such design is for simplicity, and we will discuss the possible extensions in section III. At the reader side, the collision free bits of the first received packet (or RN) can be decoded by traditional demodulation, and we subtract a stream of all “0” bits. For the second RN, we repeat this process and finally get two “1”-“0” bits. Moreover, the sample offset and amplitude of two collided signals can be easily calculated, so that we can know which RN each “1”-“0” bit belongs to.

B. The Benefits of DDC

To see the benefits of DDC, we first introduce the EPC C1G2 RFID protocol [9]. In a typical RFID system, the reader acts as a power supplier and operator of the passive tags. The reader transmits Continuous Wave (CW) to power up the tags and initiate Query Rounds to identify the tags. An illustration of tag identification is shown in Fig. 2. In each Query Round,

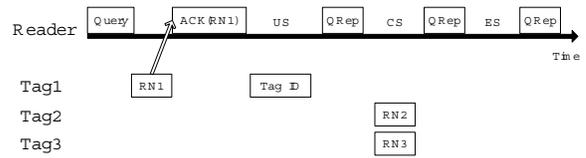


Fig. 2. EPC C1G2 UHF RFID Protocol. The RFID uses Query command to start a Query Round and uses Query Repeat(QRep) command to divide the time into slots. Each tag uses the generated random number to hash the frame length for randomly slot selection. A time slot can be Useful Slot(US),Empty Slot(ES) or Collision Slot(CS). The slot is useful only when one tag responds in that time slot.

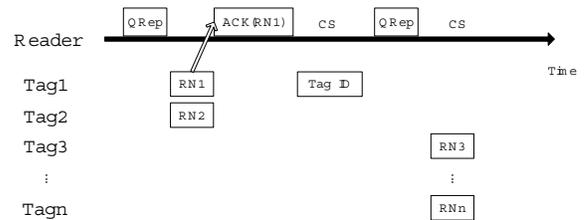


Fig. 3. The Benefits of DDC. It uses the old RN for slot selection with no modification and generates new designed RN (RN1,RN2,...,RNn) to contend for identification. The gain of DDC comes from two aspects. First, it shortens the frame length to produce more Collision Slots, thus the cost of Empty Slots is reduced. Second, it makes the collided RNs decodable, therefore, CS = US.

the reader broadcasts the tags a frame length f in the Query command, which is a tunable parameter and follows $f = 2^Q$. Each tag maintains a Random Number Generator (RNG) to generate random numbers. The random number hashes the frame length to randomly pick one slot (stored into the slot counter) in the range of 0 to $2^Q - 1$, inclusively. In each time slot, the tag transmits the RN if its slot counter is equal to zero. If one RN is received and correctly decoded by the reader, the reader sends an ACK contains the random number to the corresponding tag, and then the tag sends its ID if it finds a matched random number. We call this time slot a useful slot (US) and the whole process is tag identification. If more than one tag or none response in one time slot, a collision slot (CS) or an empty slot (ES) occurs. Consequentially, the reader fails to receive a random number. The reader will send a Query Repeat Command (QRep) to enter the next time slot, and the slot counter stored in each tag will be decreased by one. The unsuccessful identified tags will contend for identification in the next Query Round.

The proportion of useful slots is 36.8% at most in such a ALOHA based protocol. DDC is designed to reuse the collision slots as the single response slots to breakthrough this limitation. DDC is a leveraging collision design, even there is a collision in a slot, the RNs sent by the tags can still be decoded, and one of the collided tag can still achieve identification opportunity. Therefore, the efficiency can be dramatically improved. As tag use RN to contend for channel access, the collisions in RFID systems are caused by the RNs. We redesign the data pattern of RN, under which two RNs collided with each other can be decoded correctly. As

shown in Fig. 3, DDC bring two benefits for RFID systems on tag identification. First, we can correspondingly minimize the frame length to reduce the number of empty slots, which can not be achieved by traditional RFID system, as short frame length brings more collisions. Second, the approach to decoding the collision make the collision slots are useful, as DDC can decode the collided RNs.

DDC can also facilitate tag number estimation, which refers to counting the number of tags. Current schemes estimate the number of tags by the probability of empty slot or collision slot occur, a high probability of empty slot occur means a small size of tag population. One usually run the estimation many times to tell the number of tags more precisely [1], [3], [6]. Current algorithms can only use the information about collision slot, empty slot or single response slot. DDC provides information about how many tags response in one slot, which bring new opportunity and challenge for new tag number estimation algorithms. We do not provide any well designed tag estimation algorithm in this paper, but focus on the practicability study of DDC. We left the detail tag estimation algorithms for the future work.

III. THE COMPONENTS OF DDC

This section presents the detail about how to directly decode the collisions, including the amplitude and phase estimation of collided RNs etc. The discussion here involves two collided RNs, but we consider it as an iterative process and can be expanded to the case of three or more collided RNs. The main components of DDC, including collision detection, costas loop, construction and subtraction.

A. Collision Detection

The traditional method for recognizing a packet or collision is energy detection. When there is an energy jump, we can tell there is a packet appear or a collision happen to a packet. This is not practical for collision detection as the collision may increase or decrease the energy level, which due to the unpredictable phase difference between the collided packets. Notice that the RNs are fixed length stream, so that a shifted collision will produce a longer length stream. We can roughly tell there are more than one random number in the received symbols by estimating the length of collision. If we find a collision, we move the samples to next block to try to decode it.

B. Amplitude Estimation

There are always some collision free symbols in the preamble of the first received RN. The collision free symbols are also where we bootstrap from. We can estimate the amplitude of the first RN by these collision free symbols. What we received is a stream of complex number $y[t]$, and the amplitude of first collision free complex numbers can be calculated as

$$\|y[t]\| = \sqrt{I_1[t]^2 + Q_1[t]^2}. \quad (1)$$

Where I_1 and Q_1 can be precisely estimated by the collision free symbols. The amplitude of the second RN can be

estimated from the tail of the collision [8]. Here we want to present a general form of amplitude estimation to fit general cases, e.g., three collided RNs. The parts of interest is the collided segmentations, the amplitude of these parts can be represented by the following equation.

$$\|y[t]\| = \sqrt{(I_1[t] + I_2[t])^2 + (Q_1[t] + Q_2[t])^2}. \quad (2)$$

The complex stream of the first RN can be predicted, and the rest after subtraction is the second one and the amplitude of it can be easily calculated.

C. Frequency and Phase offset Estimation

Imagine one carrier wave, say $\cos(2\pi f_c t + \theta[t])$, transmitted from the tag to the reader, there are always some offsets on frequency and phase between the transmitted and received carrier waves, which due to the imperfect hardware, e.g., the oscillator. To present the offsets or errors mathematically, we have the following equation.

$$y[t] = A \cos(2\pi(f_c + f_{err})t + \theta[t]) + G[t]. \quad (3)$$

Typically, the receiver estimates the frequency error and compensates for it by estimation mechanisms, e.g., costas loop [19]. The frequency is the time derivative of the phase, so the estimation of phase offset is involving that of the frequency error [19]. We use linear regression to estimate the phase offsets of the collision free symbols and predict those of the symbols in the collision field, as follows.

$$\theta_{err}[t + T] = \theta_{err}[t] + \alpha * T. \quad (4)$$

D. Known Symbols Construction and Cancellation

The sample position of a signal is random, which produce the sample offset. The sample offset is difficult to estimate and make the sample value difference varies in different sample positions. Instead of estimating the sample offset, we construct a RN wave with N samples per symbol at the reader side. Note that this is a one time cost and do not need to communicate with the tag. To get a copy of the received RN, which we assume a stream of n samples per symbol, we find the best match one in the N/n possibilities. Now we get the amplitudes and phases of the first RN by the collision free bits. We can reconstruct this RN with the known symbols and best match pattern. As we are not sure about the position where the “1” bit starts, we can not subtract the original RN pattern. We replace the original “11” bits into continuous “00” bits in the data field. For the first RN, we exactly know what is left after subtraction but do not know the position of it only. We do exactly the same operations on the second RN, and finally two known patterns left.

It is very likely that the remaining symbols of the first random number are collided with the header of the second random number, and the header can be found by the correlation detection [7]. It also produces a huge phase tracking error, where we can identify and ignore these symbols when constructing the second random number. Here we leverage

the intrinsic property of BPSK, where the imaginary parts of the samples should be zero without collision, so we can use this property to find the dirty symbols and ignore them when tracking the phase errors.

E. Obtaining the RNs

If the remaining symbols are not overlapped, the traditional method can simply decode them and record the start positions of them. The left bits of the two RNs are likely collided at the same position in the time domain, we solve this problem by identifying the start position of this collision and match it with the time shifts of these two RNs. We can decode them even when the two “1”-“0” bits are entirely overlapped, as the sample distances to the start point of these two RNs can still be calculated. This is a special case and just suitable for decoding two collided RNs. Meanwhile, we will drop the entire collision if there is no header left but the noise level exceed a threshold.

IV. EXPERIMENTAL EVALUATION

In this section, we study the performance of our approach use results from the software defined radios testbed. There are 8 tags and 1 reader in our evaluation. Though the most application scenarios need hundreds or even thousands of tags, we claim that our evaluation is equivalent to evaluating a large scale RFID system. One can adjust the frame length to fit the number of tags, so that it is possible to produce the same percentage of useful slots in any scale number of tags. We will show that DDC can outperform the current protocol by the same parameter (frame length) setting in any scenarios. We focus on the performance improvement on tag identification in this evaluation.

A. Collision Decoding Performance

For SNRs in the range from 4dB to 18dB, we run the collision detector of DDC on sets of 1000 collisions and 1000 collision free RNs. The average false positive rate (single RN mistaken as collision) is 0% and the average false negative rate (collided RNs mistaken as clear packet) is 0.5%. To investigate the performance of DDC under different SNR levels, we run DDC for the 1000 collided RNs (each collision has exactly two RNs). The random numbers are pre assigned and we do not evaluate the impact of random number confliction at this stage (Notice that the probability of confliction is very small).

First of all, we would like to understand the impact of the signal to noise ratio (SNR) and signal to interference and noise ratio (SINR) on DDC’s performance. For the comparable SNR (+/-1dB) levels of two collided RNs, the decoding ratio is curved with respect to the SNR. The results are shown in Fig. 4. The word “one” means the decoded RNs are at least one is right. Notice that our purpose is that extract one RN from the collision. The word “both” means both RNs are successfully decoded.

Besides the impact of SNR, we also illustrate how SINR influence DDC’s performance in Fig.5. We fix the power of one RN and adjust another’s. A higher SINR make the weak one difficult to be decoded, as the subtraction process will

introduce a high noise to the rest one. And it works well if the lower SNR RN arrives first. We do not distinguish this difference and only take the average of decoding ratio in this evaluation. The EPC can decode some of the RNs as the data pattern is previously known and the one with higher energy is easy to decode.

We conduct the similar experiment for three collided RNs and the results are shown in Fig.6. The decoding process is an iterative process, The decoding ratio of three collided RNs is a little lower than that of two collided RNs , as more collided RNs will introduce more noise to the rest signals. Our empirical study show that our design works well in three collided RNs at this stage. Similarly, The word “one” means the at least one RN is successfully decoded. The word “three” means all RNs are successfully decoded.

This experiment reveals the following conclusions:

- DDC is capable of multi-collided RNs decoding, as it is an iterative decoding process. It can reuse the collision slot and makes no overhead in empty or single tag response slot.
- DDC can outperform EPC in different SINR level. DDC can successfully decode the RNs even in high SINR. EPC can only possibly decode the RN with high SNR, which is known as capture effect [22].

V. RELATED WORK

Related work falls in the following two areas.

A. Tag Identification

Tag Identification is an important issue in RFID systems. The tag identification algorithms can be classified into two categories, tree based algorithm [14], [17] and ALOHA based algorithm [2], [4], [9], [15], [16]. The tree based protocol takes the advantage of tag ID that matches query string with the prefix of it. The performance of such a protocol depends on the distribution of tags’ IDs. A smart trend traverse (STT) protocol [17] is proposed to tolerate different ID distributions. In general, tree based algorithms are still inefficient for RFID systems as the reader need to transmit a lot of commands to traverse all the possible IDs. Therefore, instead of tree based algorithms, today’s commercial RFID reader uses ALOHA based algorithm, which is our focus.

ALOHA [9], [12] is a well known MAC layer protocol to prevent collisions caused by concurrent transmissions. However, due to its pure random nature, it has a low throughput. RTS/CTS is designed [13] to address hidden terminal problem but known as high cost on channel access contention. ALOHA is still popular for its simplicity in RFID and 802.15 systems [9], [12]. Today’s ALOHA uses a contention scheme similar to RTS/CTS to reduce the cost of collisions slots, where the tag transmits a random number to contend for channel access in randomly picked time slot, because the collisions of RNs cause less cost than that of tags’ IDs. The system is still inefficient as the total cost of empty slots and collisions slots is still very high. The throughput of ALOHA system can be optimized by setting the frame length to be equal to the number of tags,

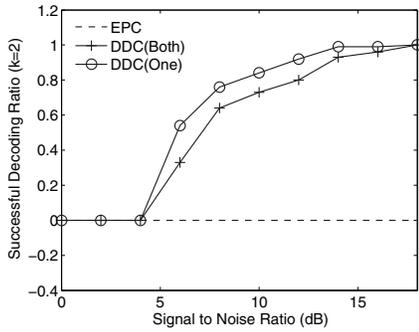


Fig. 4. Decoding Two Collided RNs

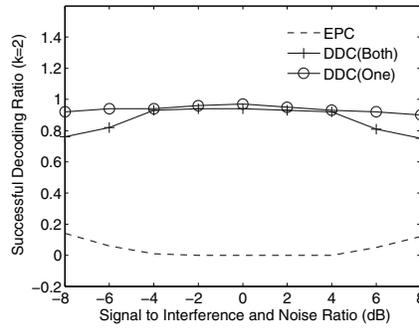


Fig. 5. Decoding Two Collided RNs for Different SINRs

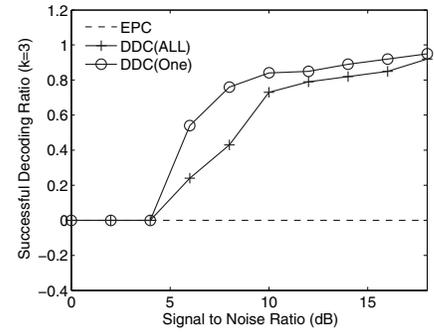


Fig. 6. Decoding Three Collided RNs

various methods [4], [9], [15], [16] are designed to adjust the frame length to maximize the throughput. A probabilistic model is proposed to enhance the RFID system performance in a mobile environment [4]. M. Buettner and D. Wetherall [2] propose tuning the physical layer operating parameters to increase the tag read rate. All the current work focus on avoid the collision, but we adopt a new approach that can decode the collided RNs. In this way, we can reuse the collision slots to enhance system performance.

B. Exploiting Collisions in Wireless Networks

Previous works which are able to decode the collisions only when multiple collisions of the same packets are available [7] or when a preknowledge of one packet is known already [8]. we are able to decode in one collision without the knowledge of any packet. Another recent work on Side Channel design [21] leverages the interference special designed interference patterns to build a free in-band Side Channel, which will degrade the Main Channel's resilience to collision. Our approach does not influence the data (tag ID) transmission in RFID systems.

VI. CONCLUSION

In this paper, we propose a new scheme that is able to directly decode the collisions (DDC). It enables the RFID reader to decode collided RNs to leverage the collision slots by preassigned RN pattern. The principle is that DDC reduce the information carried by each RN to enable concurrent transmission of multiple tags. Making use of the collision slots and reducing the empty slots, DDC enables RFID system to be able to identify the same number of tags using smaller frame length compared with traditional system. Both theoretical analysis and experiment results shows that DDC enhances the tag read rate by roughly 40% compared to the traditional protocol.

VII. ACKNOWLEDGEMENT

This research was supported in part by Hong Kong RGC Grant HKUST 617710, China NSFC Grants 60933011 and 60933012, and the Science and Technology Planning Project of Guangdong Province, China under Grant No. 2009A080207002.

REFERENCES

- [1] M. Kodialam and T. Nandagopal, "Fast and Reliable Estimation Schemes in RFID Systems", *In ACM MobiCom 2006*.
- [2] M. Buettner and D. Wetherall, "An Empirical Study of UHF RFID Performance", *In ACM MobiCom 2008*.
- [3] C. Qian, H.-L. Ngan, and Y. Liu, "Cardinality Estimation for Large-scale RFID Systems", *in IEEE PerCom 2008*.
- [4] L. Xie, B. Sheng, C. C. Tan, H. Han, Q. Li and D. Chen, "Efficient Tag Identification in Mobile RFID Systems". *In IEEE Infocom 2010*.
- [5] H. Han, B. Sheng, C. C. Tan, Q. Li, W. Mao and S. Lu, "Counting RFID Tags Efficiently and Anonymously". *In IEEE Infocom 2010*.
- [6] T. Li, S. Wu, S. Chen and M. Yang, "Energy Efficient Algorithms for the RFID Estimation Problem". *In IEEE Infocom 2010*.
- [7] S. Gollakota and D. Katabi, "Zigzag decoding: Combating hidden terminals in wireless networks". *In ACM SIGCOMM 2008*.
- [8] S. Katti, S. Gollakota and D. Katabi, "Embracing Wireless Interference: Analog Network Coding". *In ACM SIGCOMM 2007*.
- [9] EPCglobal, "EPC Radio-Frequency Identity Protocols Class 1 Generation 2 UHF RFID Protocol for Communications at 860MHz-960MHz", *version 1.2.0, 2008*.
- [10] Lei Kang and Lionel Ni, "Revisiting ALOHA with Physical Layer Network Coding". *Technical Report, CSE Dep. HKUST, 2009*.
- [11] M. Z. Brodsky and R. T. Morris, "In Defense of Wireless Carrier Sense". *In ACM SIGCOMM 2009*.
- [12] LAN/MAN Standards Committee, "Wireless Medium Access Control and Physical Layer Specifications off Low-Rate Wireless Personal Area Networks". *IEEE Std 802.15.4a-2007*.
- [13] LAN/MAN Standards Committee, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications". *IEEE Std 802.11-2007*.
- [14] V. Namboodiri and L. Gao, "Energy-Aware Tag Anti-Collision Protocols for RFID Systems", *in IEEE PerCom 2007*.
- [15] F. Schoute, "Dynamic Frame Length ALOHA", *IEEE Trans. Commun., vol. 31, no. 4, pp. 565-568, Apr. 1983*.
- [16] V. Anantharam, "The Stability Region of the Finite-user Slotted ALOHA Protocol", *IEEE Trans. Inf. Theory, vol. 37, no. 3, pp. 535 - 540, May 1991*.
- [17] L. Pan and H. Wu, "Smart Trend-Traversal: A Low Delay and Energy Tag Arbitration Protocol for Large RFID Systems", *In IEEE Infocom, 2009*.
- [18] M. Buettner and D. Wetherall, "A Flexible Software Radio Transceiver for UHF RFID Experimentation", *UW CSE Technical Report, 2009*.
- [19] J. Feigin, "Practical Costas Loop Design", *RF Design, pp. 20-36, January 2002*.
- [20] G. R. Danesfahani and T.G. Jeans, "Optimisation of modified Mueller and Muller algorithm", *Electronics Letters, Vol. 31, no. 13, 22 June 1995, pp. 1032 - 1033*.
- [21] Kaishun Wu, Haoyu Tan, Yunhui Liu, Jin Zhang, Qian Zhang, and Lionel M.Ni, "Side Channel: Bits over Interference", *In ACM MobiCom 2010*.
- [22] J. Lee, W. Kim, S.J. Lee, D.Jo, J.Ryu, T.Kwon, and Y.choi, "An Experimental Study on the Capture Effect in 802.11a Networks", *WiNTECH 2007*.