An Information-theoretic Security Metric for Future Wireless Communication Systems

Zhen Cao, Hui Deng, Lu Lu, Xiaodong Duan
China Mobile Research Institute, Beijing 100053, China
{caozhen, denghui, lulu, duanxiaodong}@chinamobile.com

Abstract—**Quantitative analysis of security properties in wireless communication systems is an important issue; it helps us get a comprehensive view of security and can be used to compare the security performance of different systems. This paper analyzes the security of future wireless communication system from an information-theoretic point of view and proposes an overall security metric. We demonstrate that the proposed metric is more reasonable than some existing metrics and it is highly sensitive to some basic parameters and helpful to do fine-grained tuning of security performance.**

I. Introduction

The emerging of novel Internet applications and services are shaping the future wireless communication systems into an open and integrated service environment. A lot of efforts are making this happen. For example, the architecture of LTE leverages an all-IP network architecture to enable mobile operators to integrate the core with the access network, providing real-time voice service and broadband IP services from the core to the mobile station. WiiSE (Wireless IP/Internet Service Environment) [1] proposed by China Mobile introduces an evolutional architecture that extends the the flat architecture of LTE-SAE with an aggregated base station and gateway entity called WiiSE Node, adding more flexibility. These types of future communication systems are envisioned to provide end users with more convenience and in the meantime making the network a flat and open environment for new innovations to take place. While these visions of future communication systems are brought into reality, they all are exposed to various security design challenges. First and foremost, scalability is a big challenge for security. Many legacy security designs do not scale well when network size increases to a certain degree. Secondly, the flat and highly distributed network architecture will move some network capabilities closer to end users, hence exposed to more risk than before. Further more, the open and all-ip environment makes it easier and more cost-efficient for the malicious to launch attacks.

Now that security is an indispensable issue for future wireless communication systems, much research and engineering efforts devote to this topic. While new architectures and designs are blooming, we still do not have an integrated way to measure and compare their performance in-depth. As a saying goes, if we cannot measure it, we cannot improve it. However, security analysis techniques are far from satisfactory in the literature. If we look at the content of the "security consideration" section or many technical documents, we are only able to get some high level understanding of security issues and risks in using the designed architecture or protocols, but how they perform overall and their impacts to the whole system are not specified quantitatively.

The lack of a proper security metric hinders security design in future wireless communication systems in many ways. First and foremost, we fail to have an integrated way to measure how those schemes perform overall, and cannot compare the performance of the proposed schemes quantitatively. Secondly, a good metric usually is able to imply a lot on how to improve the design, without which arbitrary efforts sometimes will be scattered over the plane without catching the most valuable point.

With the above in mind, this paper proposes an information-theoretic security metric for future communication systems. The notion of "Entropy" is the central theme of Information Theory. Entropy characterizes the uncertainty of random variables and also reflects the amount of information contained. How to bridge the security performance with this information-theoretic metric is an interesting problem. If we consider the underlying secure scheme as a black box in the middle that judges the validity of each input data, the security performance of the scheme can be viewed as its ability to reduce the uncertainty of the target input variable given its output. We hence define a new security metric in this way and show that it not only is helpful for administrator to do find-grained tuning of the security performance but also performs better than some existing security metrics.

II. An Abstract View of Security

Let's revisit the security of wireless communication systems (WCS) from a very abstract level. If we view the system as a functional black box in the middle, it receives input data from the signal plane and the media plane and feeds its output to the corresponding end points. Based on its security criterion, the WCS may accept the input as normal or filter out the data as intrusive. For example, authentication requests and data packets for legitimate users will pass through the security validation and be delivered to the other end point. On the contrary, requests from invalid users will be discarded. From a very abstract level, every input data has either an intrusive or normal status. As in Figure.1, we model the input of WCS as a random variable X, where $X = 1$ represents an intrusion while $X = 0$ represents normal traffic. A base rate for attacks, denoted as "B", is used to characterize the apriori probability of attacks, i.e., $P(X=1)=B, P(X=0)=1-B$. The output of the WCS is denoted as another random variable Y. Y also has two statuses. $Y=1$ means the data is discarded as intrusion or does not reach its assumed destination, while $Y=0$ means the WCS delivers this data as normal.

As in Figure.1, an intrusion input has a probability $P(Y=0|X=1)$ of being considered as normal by the WCS. This is the False Negative rate (FN) denoted as β. Cryptography schemes and authentication mechanisms are instant remedy to this attack. Even though, mis-configuration of security policy may cause illegal input accepted by the WCS. Also attackers may compromise keys from valid users and inject fake information into the network. This is generally considered as the false negative attack in our model.

Similarly, a normal event also has a probability $P(Y=1|X=0)$ of being considered as an intrusion. This is the False Positive rate (FP) denoted as α. For example, mis-configuration of routing policy may lead some packets not able to reach the correct destination, such as prefix hijack attack in BGP routing. This is generally considered as the false positive attack in our model. We will use the notations (B,α,β) throughout this paper.

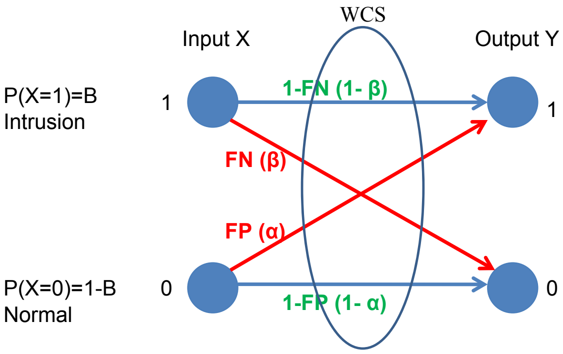


Fig. 1. Abstract Model Wireless Security

A. Need of a Good Metric

From the abstract model depicted in Figure.1, we can observe there already exist some metrics to measure the security of WCS. For example, the false negative rate and false positive rate. We may ask why these metrics are not enough for our understanding of WCS security. The answer of this question is not trivial. We would like to start with some intuitions about a well-defined metric.

First, a good metric should be comprehensive. Like we use Gross Domestic Product (GDP) to measure the economy, the metric for WCS should also be comprehensive so that it reflects the overall performance from many prospectives. Secondly, it had better be highly sensitive to some base parameters, so that fine-grained tuning of the system performance is possible. Last but not the least, a good metric should be easy to compute. This is natural because we need to find a good way to figure out the metric.

Existing metrics like the false positive rate α and false negative rate β do not fulfill the above requirements; they only reflect the performance of the system from one particular perspective, thus not comprehensive enough.

We can also use some Bayesian metrics like the Positive Predictive Value (PPV) and the Negative Predictive Value (NPV) to measure the security performance of WCS. They are defined as below.

Definition 1: The Bayesian positive detection rate (PPV) is defined as:

$$PPV = \frac{P(X=1,Y=1)}{P(Y=1)} = \frac{B(1-\beta)}{B(1-\beta)+(1-\beta)\alpha}$$

Definition 2: The Bayesian negative detection rate (NPV) is defined as:

$$NPV = \frac{P(X=0,Y=0)}{P(Y=0)} = \frac{(1-B)(1-\alpha)}{B\beta+(1-B)(1-\alpha)}$$

Clearly PPV is the probability that WCS correctly filters out intrusive data, while NPV is the probability that WCS accurately accepts normal data. In terms of usability, they are very important; the WCS is only useful only if it has high PPV and NPV. Both PPV and NPV are functions of variables (B,α,β). But each of them alone does not reflects the WCS's overall security performance. Actually PPV is not sensitive to the false negative rate β, and NPV is not sensitive to the false positive rate α. The next subsection will present the detailed analysis.

Since existing metrics do not meet our requirements for a good metric, we plan to design a new metric to measure the overall security performance from an information-theoretic point of view. Before present this metric, we first present the basic background of Information Theory.

B. Information Theory Background

Definition 3: The entropy a discrete random variable X is defined by [2]

$$H(X) = -\sum_{x\in\mathcal{A}_X} p(x)\log p(x)$$

where $\mathcal{A}_X$ denotes the set of all possible values for variable X, with the convention that $0\times\log\frac{1}{0}=0$ since $\lim_{\theta\to 0^+}\theta\log\frac{1}{\theta}=0$

The entropy of a random variable represents its "uncertainty". The more uncertain the X is, the larger the $H(X)$. $H(X)$ achieves its maximum value $\log(|X|)$ when variable X flows an even distribution among all the discrete values. If any one value of X occurs with probability 1, $H(X)=0$.

Definition 4: Let $p(x,y)$ be the joint distribution of random variables (X,Y), then the conditional entropy of X given Y is defined as [2]:

$$H(X|Y) = -\sum_{y\in\mathcal{A}_Y}\sum_{x\in\mathcal{A}_X} p(x,y)\log p(x|y)$$

The conditional entropy $H(X|Y)$ measures the average of uncertainty that remains about X when Y is known. If X is completely determined by the value of Y, $H(X|Y)=0$. On the contrary, if X and Y are statistically independent, $H(X|Y)=H(X)$, i.e., knowledge of Y does not help to determine X.

Definition 5: Consider two random variables X and Y with a joint probability mass function $p(x,y)$ and marginal probability mass functions $p(x)$ and $p(y)$. The mutual information $I(X;Y)$ is defined as [2]:

$$I(X;Y) = \sum_{x\in\mathcal{A}_X}\sum_{y\in\mathcal{A}_Y} p(x,y)\log\frac{p(x|y)}{p(x)p(y)}$$

Mutual information informs us the amount of information shared between two random variables X and Y. The relationship between the mutual information and entropy is shown as below:

Theorem 1: The relationship between mutual information and entropy:

$$I(X;Y) = I(Y;X) = H(X) - H(X|Y) = H(Y) - H(Y|X)$$

Proof of this theorem can be found in [2]. Mutual information measures the average reduction of uncertainty about X that results from learning the value of Y, or vice versa. From Theorem.1, we can obtain $0 \leq I(X;Y) \leq H(X)$. When X is completely determined by the value of Y, the uncertainty of X has been removed, i.e., $I(X;Y) = H(X)$. When X is independent of Y, the uncertainty of X is not affected even if Y is known, i.e., $I(X;Y) = 0$.

In retrospective of our abstract model of wireless communication security, the objective of WCS is to reduce the uncertainty of input variable X given the output Y. The amount of uncertainty reduction partially reflects the effectiveness of the designed security scheme. We will elaborate on this later.

C. Information-theoretic Security Metric

Mutual information defined in Def.5 characterizes how much uncertainty of random variable X is reduced after another variable Y is known. This can be used to measure the overall security performance of WCS. Ultimately the WCS is to classify the input into two categories (intrusive or normal), i.e., map the input variable X to output Y. For a fully capable WCS, knowing the output variable Y must reduce the uncertainty of the input variable X to a great extent (X and Y share much mutual information). However, for a poorly performed WCS, knowing of Y does not help us understand X, or in other words, X and Y share little mutual information.

The formula $R = \frac{I(X;Y)}{H(X)}$ reflects the ratio of uncertainty reduction of variable X given Y. R was proposed to measure the performance of Intrusion Detection System (IDS) in [3]. According to the end-2-end abstract view, WCS is quite similar to IDS in the sense that they are both designed to identify each data flow as normal or intrusive. But the metric R cannot be used to measure the security performance of WCS directly, because it does not take into account the "direction" of this reduction as explained below. Not all uncertainty reduction is good for the system. Actually the reduction of uncertainty has two directions: It can be a reduction of its correctness uncertainty, or a reduction of its incorrectness uncertainty. For example, if one WCS is fully capable of accepting each normal input and discarding each intrusive input, i.e., $y_i = x_i$, knowing of the output Y fully reduces the uncertainty of X, then $R = \frac{I(X;Y)}{H(X)} = 1$. This is the good direction. If, on the contrary, another WCS discards every truthful report and accepts every bogus report, i.e., $y_i = \bar{x_i}$. In this case Y also totally determines X, so we also have $R = 1$. But clearly this should have been the worst case in our model, and we should avoid this phenomenon when figuring out a proper metric.

Actually the direction of uncertainty reduction has been been characterized in the variable α and β. Smaller α and β represent the right direction. If we consider multiplying $\frac{I(X;Y)}{H(X)}$ with $1-\alpha$ and $1-\beta$, we can get a synthetic performance metric for the WCS as follows.

Definition 6: Let X be the random variable representing the WCS input and Y the random variable representing its output. The information-theoretic security capability C_S is defined as:

$$C_S = (1-\alpha)(1-\beta)\frac{I(X;Y)}{H(X)}$$

Note: there is a more detailed usage of C_S in [4], where false alarms in wireless sensor networks are modeled using this metric.

The computation of C_S is not difficult. Once we know the (B, α, β), we can figure out C_S easily using the following procedures.

From Def.3, we can figure out the value of $H(X)$ using Eqn(1).

$$H(X) = -B\log B - (1-B)\log(1-B) \tag{1}$$

From Def.4, we can figure out the value of $H(X|Y)$ using Eqn(2).

$$\begin{aligned} H(X|Y) &= -\sum_y\sum_x p(x,y)\log p(x|y) \\ &= -B(1-\beta)\log\frac{B(1-\beta)}{B(1-\beta)+(1-B)\alpha} \\ &- B\beta\log\frac{B\beta}{B\beta+(1-B)(1-\alpha)} \\ &- (1-B)(1-\alpha)\log\frac{(1-B)(1-\alpha)}{(1-B)(1-\alpha)+B\beta} \\ &- (1-B)\alpha\log\frac{(1-B)\alpha}{B(1-\beta)+(1-B)\alpha} \end{aligned} \tag{2}$$

Then using Theorem.1 and the above Eqn(1) and Eqn(2), we can figure out C_S easily, that is:

$$C_S = (1-\alpha)(1-\beta)\frac{I(X;Y)}{H(X)} = (1-\alpha)(1-\beta)(1-\frac{H(X|Y)}{H(X)}) \tag{3}$$

Theorem 2: The filtering capacity C_S satisfies: $0 \leq C_S \leq 1$.

This theorem is obvious, since all the multipliers in Eqn(3) $(1-\alpha)$, $(1-\beta)$ and $\frac{I(X;Y)}{H(X)}$ are within $[0,1]$.

We assert the effectiveness of the proposed metric C_S from the following three aspects.

First we show that the C_S behaves more reasonably than the existing metric R . When α or β increases, the system absolutely behaves worse. A rational metric should reflect this trend, however R fails to accomplish this. Figure.2 depicts the value of C_S and R respectively by fixing base rate at $B = 0.01$ and $\alpha = 0.9$ and ranging false negative rate β from $[0.1, 0.9]$. Clearly C_S declines smoothly when β increases. But the value of R decreases when $0.1 \leq \beta \leq 0.5$, and increases when $0.5 \leq \beta \leq 0.9$. Also, since we fix $\alpha = 0.9$ (the overall performance should be low), C_S stays at a low value under 0.05 but R almost reaches 0.34 when $\alpha = \beta = 0.9$. From these, we assert that C_S is more rational than R.

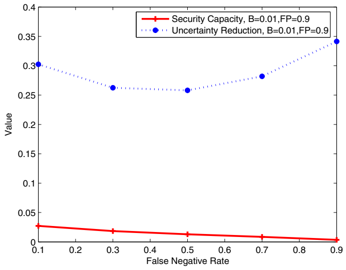
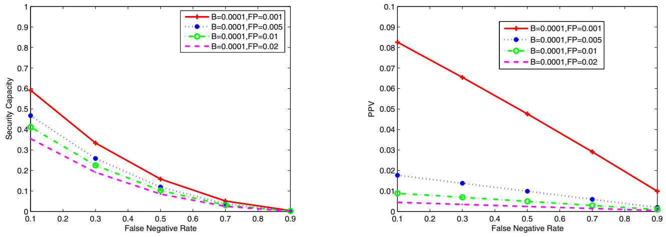


Fig. 2. Comparison with an existing metric

Secondly, Figure.3 compares the proposed metric C_S with Bayesian positive rate *PPV* in terms of their sensitivity to the false negative rate β. We fix the Base rate $B = 0.0001$ and depict four curves with the false positive rate α equals to 0.001, 0.005, 0.01, 0.05 respectively. When β ranges from 0.1 to 0.9, we plot the value of C_S and PPV in Figure.3(a) and Figure.3(b) respectively. We see that C_S poses a reasonably decline when β increases, but PPV only varies between a very small range between $[0, 1]$. This proves that C_S is more sensitive to β than *PPV*.



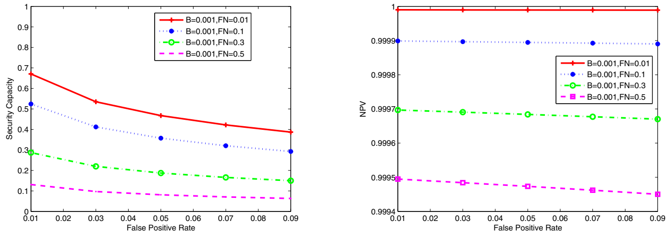
(a) C_F

(b) *PPV*

Fig. 3. Comparison of C_S and PPV, in terms of their sensitivity to False Negative Rate β

Similarly, we compare C_S with Bayesian negative rate *NPV* in terms of their sensitivity to the false positive rate α. By fixing the base rate at $B = 0.001$, Figure.4 plots four curves of C_S and NPV respectively when α ranges from 0.01 to 0.09. Clearly, when α increases, C_S decreases evidently. However, NPV is not able to reflect changes in α; it stays stably between $[0.999, 1]$. As a result, C_S is more sensitive to the false positive rate α than *NPV*.

From the above analysis, we show that our proposed metric C_S is more sensitive to the basic parameter α and β than NPV and PPV respectively. The sensitivity to the basic parameters is very important for a good metric since it is helpful to do fine-grained tuning of the system's overall performance. C_S characterizes the intrinsic performance of a wireless security scheme from an information-theoretic view. It is not only valuable for us to measure how a WCS performs overall under security breaks, but also gives us hints on how to improve its performance.



(a) C_S

(b) *NPV*

Fig. 4. Comparison of C_S and NPV, in terms of their sensitivity to False Positive Rate α

III. Conclusion

Security measurement of wireless networks is an important issue. We have raised the question of quantitative security analysis in this paper. To this end, we have proposed an information-theoretic security metric for future wireless communication systems. This metric reflects the overall security performance from many perspectives. We have demonstrated that it is more reasonable than some existing metrics and it is highly sensitive to some basic parameters and helpful to do fine-grained tuning of security performance. We believe that quantitative security analysis is an indispensable component of the security evaluation framework and also represents an effective way to improve the security design in wireless physical and higher layers.

References

[1] Zhigang Yan, Lei Lei, and Mo Chen, "Wiise - a completely flat and distributed architecture for future wireless communication systems," in *WWRF 21*, 2007.

[2] David J.C. MacKay, *A Short Course in Information Theory*, http://www.inference.phy.cam.ac.uk/mackay/info-theory/course.html, 1995.

[3] Guofei Gu, Prahlad Fogla, David Dagon, Wenke Lee, and Boris Skoric, "Measuring intrusion detection capability: an information-theoretic approach," in *ASIACCS*, 2006, pp. 90–101.

[4] Zhen Cao, Hui Deng, Zhi Guan, and Zhong Chen, "Information-theoretic modeling of false data filtering schemes in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 8, no. 2, 2012.