Analysis of Alice-Bob-Eve scenarios for secret key generation from random channels

Alain Sibille
[2]Telecom-ParisTech, 46 rue Barrault, 75013 Paris Cedex 13, France.
E-mail: alain.sibille@telecom-paristech.fr

Abstract

We address fundamental issues related in providing randomness and security for secret key generation from radio channel characteristics, based on an analysis of the knowledge shared by Alice/Bob/Eve.

1. Introduction

We here address the fundamental meaning of secret key generation (SKG) from channel randomness, in the context of the Alice/Bob/Eve trio, where Alice and Bob are legitimate users and Eve the eavsdropper. Basically, the idea of SKG stems from the fact that Alice/Bob share a common information (without yet specifying whether it is a scalar or a vector made of discrete, real or complex numbers), while Eve does not directly have access to this information. The secrecy comes from the fact that this information is assumed to be part of a statistical set and that Eve doesn't know which element of the set is shared by Alice and Bob. Then, clearly, the degree of secrecy is dependent on two related issues:

- what is the concerned statistical set ?
- to which part of this set can Eve limit her search of the key ?

These questions have many ramifications according to the assumptions we consider for the Alice/Bob/Eve trio and are discussed in the following sections.

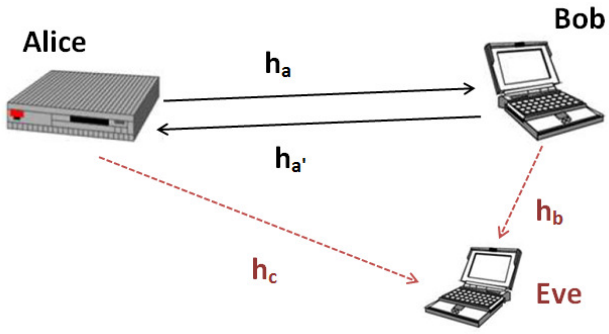

Fig. 1: Wireless communications scenario for the Alice-Bob-Eve trio

.

2. Available and secure key bits from mutual information formulas

In Information-theoretic security, the maximum number of information bits I_K that can be extracted from the estimated legitimate fading channels h_a and $h_{a'}$ (Fig. 1) can be defined as the mutual information between these channels:

$$I_K = I(h_a\,;h_{a'}) \tag{1}$$

This value corresponds to the secret key bits only in case the eavesdropper has negligible information about the legitimate channel, i.e. she is sufficiently far from both Alice and Bob. However, in principle, these bits are not fully - or not all - secure since Eve may have access to some insight about the legitimate channel. Therefore the secret key bits I_{SK} can be obtained by evaluating the mutual information between the channels seen by Alice and Bob, given Eve's observations [1-3]:

$$\mathrm{I_{SK}} = \mathrm{I(h_a\,;h_{a'}/h_b,h_c)} \tag{2}$$

3. Power/SNR aspects

Let us take a simple example: Alice is a far distant base station while Bob and Eve are in a room. We consider a single frequency Alice/Bob channel, given by a complex channel coefficient. In this case it is relatively easy to show that I_K is determined by the SNR, as given by formula below [4]:

$$\mathrm{I_K} = \log_2(1 + \frac{\mathrm{SNR}^2}{1+2\ \mathrm{SNR}}) \quad (3)$$

This value expresses that the channel experiences Rayleigh fading, in other words the real and imaginary parts are i.i.d. Gaussian distributed and the noise is additive and also Gaussian distributed and i.i.d. for the real and imaginary parts. This actually defines the statistical set mentioned above. The number of available key bits just comes from the reliability of the analog to digital conversion, which is limited by the SNR, assumed to be identical at Alice and Bob sides. Now, what about Eve ? Consider, for example, that I_K as determined above equals 10. This means that Eve has to search the key among 2^{10} possibilities. However saying that implicitly makes several assumptions:

1. the SNR is known, itself implying that the mean power and the noise power are both known from Eve
2. apart from the previous assumption, Eve has no information at all about the Alice-Bob channel coefficient

Assumption 1 is not obvious: the mean power received by Bob from Alice is not perfectly known. This mean power relates to the Rayleigh distribution, which is achieved over a limited set of realizations for Bob, e.g. when Bob moves over a "small scale" spatial area in the room. Since Eve is located elsewhere, there is no guarantee about the power estimation by Eve of Bob's received power. In addition, Eve neither knows Alice/Bob's noise powers nor Alice transmit power. So what can do Eve to guess the key ? She must make assumptions. As a result, the statistical set is larger than initially contemplated. Incidentally, this also means that the statistical distribution is no more Rayleigh.

4. Alice/Bob - Eve spatial relation

From the knowledge of the multipath structure at Alice/Bob, Eve is able to reconstruct the channel coefficients of a MIMO system. Then the problem for Eve is to determine this multipath structure, while not being collocated with Alice/Bob. The question is: how far is the multipath structure seen by Eve different from that seen by Alice/Bob ? Let us assume Eve to be in "proximity" to Bob. We can distinguish a few cases, of increasing difficulty for Eve:

3. identical multipaths, only the phase can differ, by a small amount
4. identical multipaths, the phase of the various paths differ by a significantly large amount
5. qualitatively identical multipaths, but their directions, delays and amplitudes differ
6. qualitatively different multipaths, in terms of directions, amplitudes etc.

Cases 1.-2. occur when Eve and Bob belong to the same "stationarity region". If Eve is able to obtain the parameters of the set of significant paths, she can reconstruct the signals seen by Bob. However what about Alice's channel ? Let us consider the MIMO case for the highest level of generality. The signal seen by Bob can be reconstructed by Eve for a given transmitting radiator of Alice. For another radiator, the process should be similarly repeated, since Eve is most often far or very far from Alice and cannot get any information on the channel seen by her[1]. Cases 3.-4. are much more difficult to deal with by Eve, her own estimation of Bob's channel being very uncertain, e.g. owing to obstructions, different scatterers invisibility etc. In [5], a simple model based on a disc of scatterers is used, showing the medium distance dependence of I_K/I_{SK} . Still, one very important issue for the channel reconstruction in cases 1.-2. is the knowledge of both Eve and Bob's relative positions. While Eve can possibly have an estimation of this quantity, it will be difficult for her to have a precise (sub-wavelength) knowledge, making the reconstruction process imperfect [6]. This consideration opens to a general discussion on the degree of knowledge by Eve of her own situation and that of Bob.

5. Key secrecy from Eve's point of view

Assuming a given SKG scheme, Eve's job for is to make use of the knowledge of the scheme. However, if the number of available key bits for Alice/Bob is large, this job might be extremely resource consuming, e.g. through trying all possible keys. Then Eve will try to limit this search, e.g. by reducing the size of the statistical set of possible channel coefficients. In the example of section 2 above, a single Rayleigh distributed channel coefficient was considered as the source of randomness, which is a rather moderate complexity case. In a more realistic approach, a larger number of degrees of freedom will be exploited. For instance in OFDM, we typically expect as many degrees of freedom as the number of decorrelated subcarriers. In a wide band single carrier physical layer scheme, the various multipaths will provide as many degrees of freedom [7]. The more information Eve can gather about this multipath structure, the better it is for her. For that

[1] Conceptually we can think of a spy near Bob and another one near Alice, who could exchange information about the channels seen by both legitimate users

purpose, let us generalize this knowledge acquisition by Eve. While cases 1.-2. in section 4 are based on Eve's measurement of the multipath structure to reconstruct Bob's channel, this is not a requirement if other ways to acquire information are possible. In a completely hypothetical situation, Eve would be able to acquire the exact knowledge of the radio environment (position and nature of all scatterers, of building maps, size, position and material nature of walls etc.) and she would have at her disposal a perfectly accurate propagation computing tool. If she is also able to precisely locate each Alice/Bob/eve device, then she is able to compute the channel shared by any pair of devices, whatever their location inside or outside a stationarity region. This hypothetical scenario shows that the capacity for Eve to determine the Alice/Bob channel fully depends on the knowledge acquired by Eve on the trio and on the propagation environment. In sophisticated schemes this may come from any source of information: a camera picking the scene, a GPS, a side information provide by an external tier... From this knowledge, Eve can reduce the size of the statistical set of possible channel coefficients and make her job easier.

6. Key secrecy from Bob's point of view

We identified the knowledge obtained by Eve as a major ingredient of the fundamental easiness or difficulty in obtaining the secret key. From Bob' point of view, his knowledge of Eve's situation is equally important. Indeed, assume Bob has strictly no information about Eve's location, this means Eve can be km or just cm away from Bob's antenna. Ensuring security for Bob implies to ensure it in the worst case, obviously nearly impossible. If, on the other hand, Bob is sure that Eve can only be outside a room, his confidence in the security of a key generated from his own channel will be much higher. In other words the security cannot be universally guaranteed for Bob, it depends on assumptions on Eve's situation.

7. Available key bits in relation with the scenario

Assume Bob can be at any distance from Alice, including very far ones for which the received signal vanishes. It is rather intuitive that if we use formula (1), we will basically find $I_K \sim 0$. However this result is irrelevant, since obviously if the received signal is extremely weak, Alice/Bob will never use it to generate a key. In any communication establishment protocol there is an estimation phase allowing to determine at least the RSSI, according to which Alice/Bob will agree on a suitable number of key bits. Then, I_K is the mutual information of (1), *conditional to the SNR*. The relevant statistical set is the set of channel coefficients respecting this condition, typically a stationarity region. We can generalize this reasoning: the above result $I_K \sim 0$ applies to an unrealistic situation in the absence of any a-priori information. More realistically:

- for a single antenna link at a single frequency, Alice/Bob appreciates the link quality (SNR), thus the SNR is the unique information and the considered statistical set should contain all relevant scenarios compatible with this knowledge. Typically, the channel can be either Rayleigh or Rice, then such channels should be realized in the statistical set with the appropriate statistical distribution of the K factor for the environments considered. This leads to a *single* value of I_K. If the transmitter/receiver pair is able to determine the K factor through an adequate measurement, then I_K can be computed conditionally to the K factor. This leads to a *distribution* of values of I_K
- for a multi-antenna link, Alice/Bob may be able to determine spatial/angular characteristics of the channel. Then, I_K is *conditional* to these characteristics (e.g. the direction of the BST), again leading to a distribution of I_K values

The conclusion is that, depending on the information Alice/Bob can access and on the algorithm to fabricate a key, we compute I_K conditionally to this information, resulting in a distribution of I_K values. For instance in the first item above, the distribution of Rice K factors, directly translated into a distribution of I_K values, will be chosen from what is known in the litterature for the K factors in the environments of interest (urban, sub-urban, rural...) and the proportion of these various environments represented in the chosen scenario. This makes I_K scenario-dependent, which can be used to ensure that the SKG strategy will be optimally chosen with respect to the target application scenario.

8. Secure key bits in relation with the scenario

The problem gets more tricky when incorporating Eve in the scenario and depends on the point of view:

- according to Bob's point of view to ensure maximum security, we should allow for Eve a certain set of statistical realizations. The value of I_{SK} will depend on this set, being intuitively smaller if Eve is allowed to come closer to Bob. In other words, the computed value of I_{SK} is conditional to Eve scenario vs. Bob
- according to Eve's point of view, where the goal is to attack Bob' security as much as possible, the computed value of I_{SK} is conditional to what is possible in terms of proximity to Bob

These two kinds of point of view/conditionality can therefore be determined according to the relevant scenarios/applications. The philosophy is similar to section 7, assuming environments and conditional parameters for Eve. In general, Eve can be assumed to be close to either Alice or Bob, otherwise the chance for her to capture channel data relevant to the Alice/Bob link is virtually zero. Hence the statistical set relevant to Eve involves the same type of local radio environment as Alice or Bob. However we need to consider, in addition to the value of stochastic parameters for Alice/Bob,

the corresponding ones for Eve. The complication also comes from the fact that Eve may not be instrumented like e.g. Bob but may own better SNR electronics, more antennas, directional antennas etc. Obviously it's impossible to cover all possible cases, however trends vs. the type and performance of Eve's equipment vs. Alice/Bob's may be obtainable. Again, the value of I_{SK} or the distribution of values will depend on the degree of information achieved by the trio and the fact we choose Bob's or Eve's point of view. Let us give some examples:

- Bob is at a given location in a stationarity region of a local Rayleigh distributed channel coefficients with a well defined SNR and a single antenna. He has no a-priori information on Eve but he can consider two scenarios: Eve is in the room up to 50 cm distance from him, or Eve is outside the room. In the former case, the statistical set for the channel coefficients of Alice/Bob should just be a Rayleigh distribution and a foxed SNR (see formula (3) for I_K) and for Eve it should cover all possible positions of Eve from the distance 50 cm until infinite.

In this example we will find a single value of both I_K and I_{SK} . There is a good chance that $I_{SK} = I_K$. Indeed, If Eve can be at any distance from Bob between 50 cm and infinity and if we don't weight the probability of his position, the relative probability for Eve to be very close to Bob will be zero. This is nonsense with respect to our problem, because a spy will get as close as possible to Bob. In other words, the scenario to consider for Eve should be realistic: we should define a set of relative positions for Eve from the minimum realistic to the maximum realistic, eventually putting probability weights to Eve/Bob distance. Then the quality of the SKG process will be related to the scenario.

9. Main conclusions on scenarios vs. SKG from channel randomness

The discussion developed in this paper has attempted to clarify the fundamental aspects in the security of key generated from random channels. The degree of security id dependent on the degrees of freedom involved in this randomness (the more, the merrier), but also on the information that can be acquired by eve in order to reduce her lack of knowledge on Alice/Bob channel and her effort in searching among possible keys. In order to evaluate the number of available key bits for Alice/Bob to generate a secret key and the number of secure bits among the available ones, we need to compute mutual information quantities, which are fundamentally expressed in terms of a statistical ensemble for the Alice/Bob/Eve trio. This statistical ensemble stems from the choice of the scenario, itself related to the assumptions considered in a practical application - or use case - for the trio. There are a large number of potential parameters, environments, situations for Alice/Bob/Eve that are part of these assumptions or choices. Therefore, the evaluation of the available and of the secure bits will be valid in relation to these choices, in other words they are not absolute but specific to the scenario.

Fortunately, the potential richness of the radio channel in its various components bring hopes for adequate randomness and degrees of freedom, for realistic cases in terms of security. It is possible to exploit the propagation channel richness in the delay or frequency domain, its temporal variability, its angular variability through multiple antenna techniques, and its antenna variability

10. Acknowledgments

Part of this work has been funded by the European Commission through the FP7 project PHYLAWS (grant agreement n° 317562).

11 References

1. M.Bloch and J.Barros, "Physical layer security - from information theory to security engineering," Cambridge University Press, 2011.
2. U.Maurer, "Secret key agreement by public discussion from common information," IEEE Transactions on Information Theory, 1993, pp. 733-742.
3. J. Wallace and R. Sharma, "Automatic secret keys from reciprocal MIMO wireless channels: measurement and analysis," IEEE Trans. Inf. Forensics and Security, vol. 5, no. 3, pp. 381-392, Sep. 2010.
4. F. Amoroso and W. W. Jones, "Geometric model for DSPN satellite reception in the dense scatterer mobile environment," IEEE Trans. Commun., vol. 41, pp. 450-453, Mar. 1993.
5. Taghrid Mazloum, Francesco Mani and Alain Sibille, "A Disc of Scatterers Based Radio Channel Model for Secure Key Generation", EUCAP, The Hague, Netherlands, April 6-11, 2014
6. C. Chen and M. Jensen, "Secret key establishment using temporally and spatially correlated wireless channel coefficients," IEEE Trans. on Mobile Computing, vol. 10, no. 2, pp. 205-215, Feb. 2011.
7. R. Wilson, D. Tse and R.A. Scholtz, "Channel identification: Secret sharing using reciprocity in ultra wideband channels," IEEE Trans. Inf. Forensics and Security, vol. 2, no. 3, pp. 364-375, Sep. 2007.