DIGITAL SIGNAL PROCESSING **6,** 207–212 (1996)
ARTICLE NO. 0023

Cryptographic Key Agreement for Mobile Radio

Amer A. Hassan,* Wayne E. Stark,† John E. Hershey,‡
and Sandeep Chennakeshu*

*Ericsson Mobile Communications Inc., 1 Triangle Drive, Research Triangle Park, North Carolina 27709; †EECS Department, The University of Michigan, Ann Arbor, Michigan 48109; and GE Corporate R&D, 1 River Road, Schenectady, New York 12301

Hassan, A. A., Stark, W. E., Hershey, J. E., and Chennakeshu, S., Cryptographic Key Agreement for Mobile Radio, *Digital Signal Processing* **6** (1996), 207–212.

The problem of establishing a mutually held secret cryptographic key using a radio channel is addressed. The performance of a particular key distribution system is evaluated for a practical mobile radio communications system. The performance measure taken is probabilistic, and different from the Shannon measure of perfect secrecy. In particular, it is shown that by using a channel decoder, the probability of two users establishing a secret key is close to one, while the probability of an adversary generating the same key is close to zero. The number of possible keys is large enough that exhaustive search is impractical. © 1996 Academic Press, Inc.

I. INTRODUCTION

The need for secure data communications in mobile systems is apparent. A broker or a banker might need to relay sensitive data to a selected customer through the radio channel. In one-key cryptographic techniques, such as DES [1], a common secret key is needed between such two users. Key distribution and protection is an extremely important aspect of secure communications systems. However, it is not unlikely that the two users do not share a common secret key *a priori*, but need to use a secure channel. In the case where two users do not share a common key it is possible to use two-key public key cryptography (PKC) such as described in [2].

As an exploration of an alternative to PKC, a technique proposed in [3] demonstrated that characteristics of the radio channel might be used to the same effect. These characteristics are the confirmed short term reciprocity of the channel and the rapid spatial decorrelation of phase in the radio channel; that is, for an incrementally short period of time, the impulse response of a channel viewed from the antenna from A to the antenna for B is the same as the impulse response of the channel viewed from B to A, excluding thermal noise. In this paper the performance of the key management technique that was proposed in [3] is evaluated for the purpose of practical implementation.

In addition to providing another paradigm for establishing a mutually held secret variable over a public channel, the method addressed in this paper is further motivated by the following consideration. The popular PKC algorithms are unprovably secure. The complexity of computation for conventional PKC schemes and the amount of information that must be exchanged can be quite severe. As new attacks against the PKC systems are uncovered, these systems have retreated to ever longer exchange vectors and ever more complex computations. The technique addressed in this paper is a practical alternate mechanism for cryptovariable establishment and sharing. The mechanism for secrecy in this scheme depends upon a physical process, and the cryptovariable can be established with computations equivalent to a bounded distance decoding algorithm. Thus, the decoder we use to establish the cryptographic variable may also be used for processing the subsequent data transmission. Also a conventional PKC system requires a (pseudo) random quantity to be generated by each party. With the proposed system there is no need to do this, as the randomness is provided by the non-time stationarity and non-spatial stationarity of the channel itself.

Wyner [5] and Ozarow and Wyner [6] exploit the channel for secret communications but in a com-

1051-2004/96 $18.00

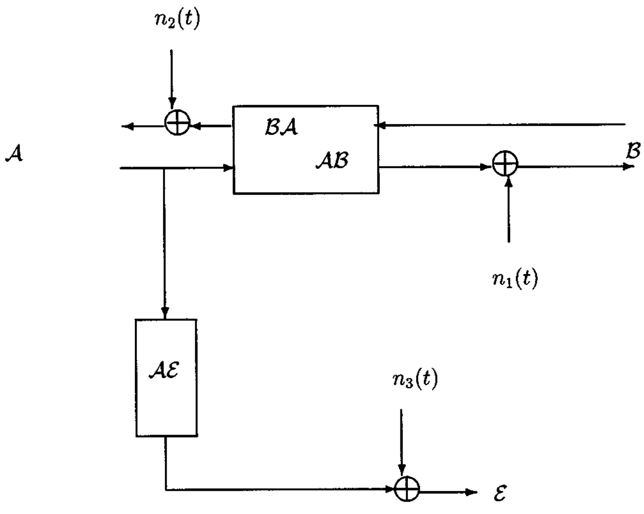

FIG. 1. A block diagram showing the system considered with two different time varying channels; $n_i(t)$, $i = 1, 2, 3$, represent thermal noise. The impulse response of channel AB is the same as the impulse response of channel BA, but it is different from that of channel AE or BE.

pletely different setting. Their results are existence proofs; in particular, it was shown that for given noisy channels (one for the communicators and one for an eavesdropper) there *exists* a rate below which *perfect* secrecy is possible. This assumes the eavesdropper has a noisier channel than the communicators' channel. In [7] the author has an interesting overview of cryptographic key agreement schemes that use the channel in different ways. In most of these schemes the two parties initially share a short secret key; the work in [7], however, relaxes this assumption and addresses the problem of generating a shared secret key by two users each knowing a random variable and the two random variables being dependent. A protocol is described to establish a common secret key and which uses a noisy channel. However, this protocol is completely impractical when the eavesdropper's channel is much better than the communicators'. The scheme described in this paper is computationally secure even when the eavesdropper has a substantially better channel than the communicators'. Finally, the recent work in [8] evaluates information theoretic models of secret sharing. The authors were able to determine the largest achievable key rate—called key capacity.

The performance measure taken in this paper is probabilistic, and different from the Shannon measure of perfect secrecy [4]. In particular, it is shown that by using a channel decoder, the probability of two users establishing a secret key is close to one, while the probability of an adversary generating the same key is essentially zero; this will be referred to as probabilistic secrecy. Also, the number of possible keys is large enough that exhaustive search is impractical, this will be referred to as computational secrecy.

The paper is organized as follows. The system model and the key distribution technique are addressed in Section II. The performance of the proposed technique is analyzed in Section III; there, numerical results using Golay coding are evaluated. In Section IV we draw conclusions.

II. SYSTEM MODEL

The proposed system is shown in Fig. 1, where E is an adversary. During the kth signaling interval $(kT, (k+1)T]$, A transmits $s(t)$ consisting of two sinusoids at frequencies f_1 and f_2 with equal phases and equal energy E; that is,

$$s(t) = \sqrt{\frac{2E}{T}} \cos(2\pi f_1 t + \phi) + \sqrt{\frac{2E}{T}} \cos(2\pi f_2 t + \phi). \quad (1)$$

The signal is transmitted through a time-varying faded channel, corrupted by additive white Gaussian noise $n(t)$ with double-sided power spectral density $N_o/2$. We assume that $\cos(2\pi f_1 t)$ and $\cos(2\pi f_2 t)$ are orthogonal and separated by at least the coherence bandwidth of the channel. Then the received signal $r(t)$, for $t \in (kT, (k+1)T]$, is given by

$$r(t) = \sqrt{\frac{2\Lambda_1^2(k)E}{T}} \cos(2\pi f_1 t + \Theta_1(k)) + \sqrt{\frac{2\Lambda_2^2(k)E}{T}} \cos(2\pi f_2 t + \Theta_2(k)) + n(t),$$

where the random variables $\Lambda_i(k)$, $i = 1, 2$, are independent and identically distributed random variables due to fading, with Rayleigh probability density functions

$$p_\Lambda(\lambda_i) = \begin{cases} \dfrac{\lambda_i}{\sigma^2} \exp\left(-\dfrac{\lambda_i^2}{2\sigma^2}\right), & \text{for } \lambda_i \geq 0 \\ 0 & \text{for } \lambda_i < 0 \end{cases}, \quad (2)$$

where $\sigma^2 = \mathrm{E}(\Lambda_i^2(k))$ is a characteristic of the channel (E denotes expectation with respect to p_Λ). The random variables $\Theta_1(k)$ and $\Theta_2(k)$ are mutually independent, each with a uniform probability density function over $[-\pi, \pi]$.

B differentially detects an estimate of $\Delta_k^B = \Theta_1(k) - \Theta_2(k)$ and quantizes the estimate into one of M phase values, say $Q(\cdot)$. Except in sequences, the time index k will be dropped with no ambiguity.

The differential baseband signal can be shown to be

$$\begin{aligned} U_B =& 2\Lambda_1\Lambda_2 E \exp\{j(\Theta_1 - \Theta_2)\} + \Lambda_1 N_1 + \Lambda_2 N_2^* \\ =& X_B + jY_B, \end{aligned} \tag{3}$$

where N_1 and N_2 are complex valued Gaussian random variables with zero mean and variance $2EN_o$, and * denotes conjugation. The estimated phase difference is given by $\Phi^B = \tan^{-1} Y_B/X_B$ and the quantizer output is $Q(\Phi^B)$.

Repeating the above transmission at times $k = 1, 2, \ldots, n$, B establishes the sequence

$$S_B = (Q(\Phi_1^B)\, Q(\Phi_2^B) \cdots Q(\Phi_n^B)). \tag{4}$$

Similarly, B transmits a sequence of two sinusoids at the frequencies f_1 and f_2 and with equal phases, after negligible delay among transmissions; that is, A transmits, then B, then A, and so on in an interleaved order to maintain the reciprocity assumption. For instance, consider a mobile with speed of 100 km/h and using a carrier in the 900 MHz region; with a delay of 10 μs, the distance moved by the mobile would be 0.00028 m, which is negligible compared to the wavelength 0.3 m. Thus A forms the baseband differential signal

$$\begin{aligned} U_A =& 2\Lambda_1\Lambda_2 E \exp\{j(\Theta_1 - \Theta_2)\} + \Lambda_1 V_1 + \Lambda_2 V_2^* \\ =& X_A + jY_A, \end{aligned} \tag{5}$$

where V_1 and V_2 are independent of N_1 and N_2. The estimated phase difference is $\Phi^A = \tan^{-1} Y_A/X_A$. Notice that due to the reciprocity of the channel, the only difference between U_A and U_B is the additive Gaussian noise.

Therefore, A establishes the sequence

$$S_A = (Q(\Phi_1^A)\, Q(\Phi_2^A) \cdots Q(\Phi_N^A)). \tag{6}$$

An adversary E will have the information

$$\begin{aligned} U_E =& 2\Lambda_3\Lambda_4 E \exp\{j(\Theta_3 - \Theta_4)\} + \Lambda_3 N_3 + \Lambda_4 N_4^* \\ =& X_E + jY_E, \end{aligned} \tag{7}$$

where Λ_i, $i = 1, 2, 3, 4$ are mutually independent, and the estimated phase difference is $\Phi^E = \tan^{-1} Y_E/X_E$; also, Θ_i, $i = 1, 2, 3, 4$ are mutually independent random variables.

The adversary E establishes the sequence

$$S_E = (Q(\Phi_1^E)\, Q(\Phi_2^E) \cdots Q(\Phi_N^E)). \tag{8}$$

The sequences S_A, S_B, and S_E are the input to an error correction decoder. The outputs of the decoders are the keys K_A, K_B, and K_E. Notice that there is no encoding performed at the transmitter end. The decoder essentially limits the number of possible keys to increase the reliability for key agreement.

III. PERFORMANCE ANALYSIS AND RESULTS

The following events will be used to assess the performance of a key distribution system:

$$G_i = \{\Phi^A \in R_i, \Phi^B \in R_i\}, \quad B_i = \{\Phi^A \in R_i, \Phi^E \in R_i\};$$

here R_i is the region in phase space that is mapped to symbol i.

The probability of a symbol match among A, B is

$$\begin{aligned} p_g =& \Pr\left\{\bigcup_{i=1}^{M} \Pr(G_i)\right\} \\ =& \sum_{i=1}^{M} (\Pr(\Phi^A \in R_i))^2. \end{aligned} \tag{9}$$

The probability of a symbol match among A, E is

$$p_b = \Pr\left\{\bigcup_{i=1}^{M} (\mathrm{B_i})\right\} = \frac{1}{M}. \tag{10}$$

The probability of an estimated phase in a decision region is derived in the Appendix.

Now consider the use of a linear block code with minimum Hamming distance d, dimension k, and a block length n equal to the length of the secret key needed. Then $t = \lfloor (d-1)/2 \rfloor$ is the number of errors that can be corrected by the decoder. The probability of agreement being in success is the probability of the two received vectors being in the same decoding region of a codeword. This can be shown to be

$$\begin{aligned} \Pr(K_A = K_B) = \sum_{l} A_l \sum_{j=0}^{l} \sum_{k=0}^{n-l} \sum_{m_1=0}^{l-j} \sum_{m_2=0}^{j} \sum_{m_3=0}^{k} \sum_{m_4=0}^{n-l-k} \\ \binom{n}{\beta} (1 - p_g)^{\beta} p_g^{n-\beta}, \end{aligned} \tag{11}$$

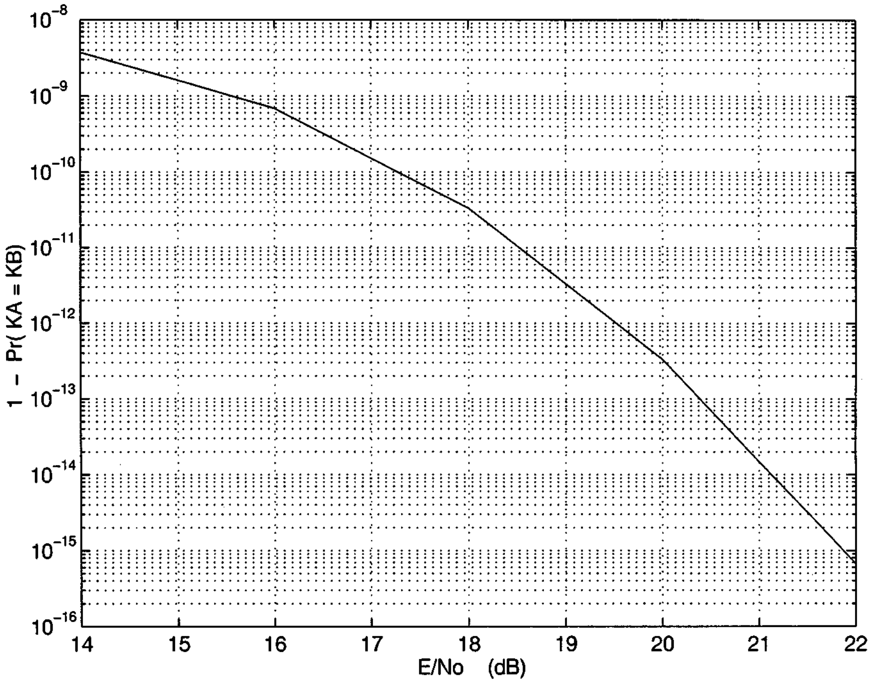

FIG. 2. The performance of (23,12) Golay code to establish a key of length 64 bits by concatenating three codewords.

where A_l is the weight enumerator function of the code, and

$$\beta = m_1 + m_2 + m_3 + m_4$$

$$0 \leq j + k \leq t$$

$$0 \leq m_1 + j + - m_2 + k - m_3 + m_4 \leq t.$$

The probability of adversary success P_B is given by a similar equation substituting p_b for p_g. For small k, more redundancy is available and a code with better error correction capability can be obtained; however, the number of possible keys becomes smaller and for small enough k exhaustive search by an adversary becomes feasible. Therefore, the choice of the code parameters are crucial, since the code restricts the key space, but the reduction should not yield an insecure system. Without the use of a decoder, $\Pr(K_A = K_B) = \Pr(S_A = S_B) = p_g^n$ and $\Pr(K_A = K_E) = \Pr(S_A = S_E) = 1/M^n$.

It is of interest to discuss more tradeoffs involved in the key distribution system. Some tradeoffs were pointed out earlier: small k yields a code with good error correcting capability, but exhaustive search becomes faster exponentially with decrease k. For large M, a larger code can be used thus increasing the computational secrecy of the system; also, p_b decreases, which results in a good probabilistic secrecy. However, this is not sufficient to obtain a good cryptographic system; with increasing M, thermal noise effects become dominant and an increase in E_b/N_o is required to achieve a key agreement with certain probabilistic secrecy. Therefore, a tradeoff exists between computational secrecy, probabilistic secrecy, and transmitted energy.

To demonstrate with an example, consider the use of a (23,12) Golay code to establish a key of length 64 bits. This key is the concatenation of three subkeys each of length 23 (the last 4 bits can be dropped). The Golay code is a perfect code, and, therefore, the decoder will always output a codeword. The probability of key establishment is then given by $\Pr(K_A = K_B)^3$, where $\Pr(K_A = K_B)$ is given in (11). Figure 2 shows the performance of the system in terms of $1 - \Pr(K_A = K_B)$ as a function of average E/N_o. The use of a decoder is necessary for A and B to establish a cryptographic key. However, a decoder does not help E.

The choice of a code and the corresponding decoder is important. One would be tempted to use a Reed–Solomon code with large M, since exhaustive search by an adversary for such codes is essentially impossible, and the performance of such codes is attractive in error control coding. The problem with using Reed–Solomon codes is the fact that the code is sparse. With very high probability the analyzed protocol will fail to establish the cryptographic key. This is true when using any sparse codes.

IV. CONCLUSION

We have evaluated the performance of an unconventional cryptographic key agreement technique based on the reversibility of a radio channel. This technique results in superior computational secrecy as well as probabilistic secrecy. Using this scheme, arbitrary long keys can be shared, and a key can change during a "session." An equivalent system would be the transmission of $2M$ orthogonal tones by each user. This system will have the same performance; however, it requires a much larger bandwidth, as required by orthogonal signaling.

APPENDIX

In this Appendix, the probability density function of Φ is evaluated. Initially, assume $\Delta = \Theta_1 - \Theta_2$ is given and equal to 0.

Consider

$$\begin{aligned} U &= 2\Lambda_1\Lambda_2 E + \Lambda_1 N_1 + \Lambda_2 N_2^* \\ &= X + jY \\ X &= 2\Lambda_1\Lambda_2 E + \mathrm{Re}(\Lambda_1 N_1 + \Lambda_2 N_2^*) \\ Y &= \mathrm{Im}(\Lambda_1 N_1 + \Lambda_2 N_2^*), \end{aligned}$$

where, conditioned on Λ_1 and Λ_2, $\mathrm{E}(X) = 2\Lambda_1\Lambda_2 E \triangleq \mu$, $\mathrm{E}(Y) = 0$, and variance (X) = variance $(Y) = 2EN_0 (\Lambda_1^2 + \Lambda_2^2) \triangleq \sigma_o^2$. The conditional joint probability density function of X and Y is

$$P(x, y|\Lambda_1, \Lambda_2) = \frac{1}{2\pi\sigma_o^2} \exp\left\{-[(x-\mu)^2 + y^2]/2\sigma_o^2\right\}$$

with the change of variables

$$R = \sqrt{X^2 + Y^2}, \quad \text{and} \quad \Phi = \tan^{-1}\frac{Y}{X}$$

the conditional joint density function of Θ and R is given by

$$\begin{aligned} &p(R, \Phi|\Lambda_1, \Lambda_2) \\ &\quad = \frac{R}{2\pi\sigma_o^2} \exp\left\{-(R^2 + \mu^2 - 2\mu R\cos\Phi)/2\sigma_o^2\right\}. \end{aligned}$$

Integrating over $R \in [0,\infty)$, it can be shown that the probability density function of Φ is given by

$$\begin{aligned} p_\Phi(\phi|\Gamma) = {} & \frac{1}{2\pi}\exp\{-\Gamma\} \\ & + \frac{1}{\sqrt{2\pi}}(\sqrt{\Gamma}\cos\phi)\exp\{-\Gamma\sin^2\phi\} \\ & \times [1 - Q(\sqrt{2\Gamma}\cos\phi)], \end{aligned}$$

where

$$\Gamma = \frac{\Lambda_1^2\Lambda_2^2}{\Lambda_1^2 + \Lambda_2^2}\frac{E}{N_o}$$

and

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\{-u^2/2\}\, du.$$

The probability density function $P_\Gamma(\gamma)$ can be shown to be

$$P_\Gamma(\gamma) = \int_0^1 \frac{\gamma}{\overline{\gamma}}\frac{1}{x^2(1-x)^2}\exp\left\{-\frac{\gamma}{\overline{\gamma}}\frac{1}{x(1-x)}\right\} dx,$$

where $\overline{\gamma} = 2\sigma^2 E/N_o$ is the average signal-to-noise ratio.

The probability density function of Φ conditioned on $\Delta \neq 0$ is then given by $p_\Phi(\phi - \Delta)$. Finally, note that Θ_1 and Θ_2 are identically distributed uniform random variables over $[-\pi, \pi)$. The sum $\Delta = \Theta_1 - \Theta_2$ can take values in $[-2\pi, 2\pi)$ according to a triangular function (the convolution of two uniform distributions). Thus to resolve the ambiguities of 2π, define the random variable

$$\Delta' = \begin{cases} \Delta - 2\pi, & \pi \leq \Delta \leq 2\pi \\ \Delta, & -\pi \leq \Delta \leq \pi \\ \Delta + 2\pi, & -2\pi \leq \Delta \leq -\pi \end{cases}.$$

It can be shown that Δ' is uniformly distributed over $[-\pi, \pi]$.

With regions given by $R_i = [-(i-1)2\pi/M, i2\pi/M)$, for $i = 1, \ldots, M$, the desired probability is given by

$$\begin{aligned} &Pr(\Phi \in R_i) \\ &\quad = \frac{1}{2\pi}\int_0^\infty \int_{-\pi}^{\pi} \int_{R_i} p_\Phi(\phi - \delta|\Gamma)P(\Gamma)\, d\phi\, d\delta\, d\Gamma. \end{aligned}$$

ACKNOWLEDGMENTS

The authors thank a reviewer for pointing out a fault in using sparse codes. They also thank Professor J. Massey for his valuable comments.

REFERENCES

1. *Data Encryption Standard*. Federal Information Processing Standards, Publication Number 46, National Bureau of Standards, 1977.
2. Rivest, R. L., Shamir, A., and Adelman, L. A method for obtaining digital signatures and public key crypto-systems. *Commun. ACM* **21,** no. 2 (1978), 120–126.
3. Hershey, J. E., Hassan, A. A., and Yarlagadda, R. Unconventional cryptographic keying variable management. *IEEE Trans. Commun.*, in press.
4. Shannon, C. E. Communication theory of secrecy systems. *Bell Syst. Tech. J.* **28** (1949), 656–715.
5. Wyner, A. D. The wire-tap channel, *Bell Syst. Tech. J.* **54** (1975), 1355–1387.
6. Ozarow, L. H., and Wyner, A. D. Wire tap channel, II. *Bell Syst. Tech. J.* **63** (1984), 2135–2157.
7. Maurer, U. M. Secret key agreement by public discussion from common information. *IEEE Trans. Inform. Theory* **39** (1993), 733–742.
8. Ahlswede, R., and Csizár, I. Common randomness in information theory and cryptography. I. Secret sharing. *IEEE Trans. Inform. Theory* **39,** 1121–1132.