Reducing String Oblivious Transfer to Universal Oblivious Transfer

Stefan Wolf

Computer Science Department
Swiss Federal Institute of Technology (ETH Zürich)
CH-8092 Zürich, Switzerland
E-mail address: wolf@inf.ethz.ch

Abstract

This paper is concerned with information-theoretic reductions of 1-out-of-2 chosen string oblivious transfer to other primitives that are as weak as possible. At Eurocrypt '97, Brassard and Crépeau presented a reduction to so-called generalized bit oblivious transfer, a primitive in which the sender inputs a pair of bits of which the receiver can choose to obtain at most one bit of *deterministic* information (e.g., one of the two bits, the XOR or AND function of the bits, and so on). It was stated as an open problem how this can be generalized to *probabilistic* information (where the receiver for instance obtains noisy versions of the bits). We show that the most optimistic answer to this question is the correct one: Whenever the so-called universal oblivious transfer is such that Bob does not obtain full information about the bits sent, then string oblivious transfer can be reduced to it in linear time, where an exponentially small failure probability must be tolerated. The new technique applied in the analysis uses specific side information provided to the receiver by an oracle and allows to simplify the argumentation.

Keywords. Cryptography, oblivious transfer, universal oblivious transfer, reductions among primitives.

1 Introduction

This paper deals with information-theoretic reductions between fundamental cryptographic primitives. Examples of such primitives are oblivious transfer, secret two-party computation, or secret-key agreement. In [8], the more

general term *information-theoretic primitive* (or *IT-primitive* for short) was introduced which includes for example also the concepts of noisy communication channels or joint randomness. The reason why efficient information-theoretic reductions among such primitives are of interest is that the primitives that can easily be achieved (or exist in reality, such as a noisy channel) are probably not the same as the ones that one is interested in (e.g., for using it as a building block of a cryptographic protocol). Many authors have studied relationships between such primitives (see for example [8] and the references given therein).

An important special case are the various versions of so-called *oblivious transfer (OT)*. The standard bit OT between two parties Alice and Bob corresponds to a binary erasure channel with erasure probability 1/2 and works as follows. The sender Alice's input is a bit b, which Bob learns with probability 1/2, whereas otherwise, Bob obtains no information about b. Alice on the other hand obtains no information about the fact whether Bob received the bit or not.

In the more important *1-out-of-2 chosen bit OT* ($\binom{2}{1}$ OT for short) Alice sends two one-bit messages to Bob, exactly one of which Bob can choose to read (remaining completely ignorant about the other one) such that Alice does not get any information about which message Bob chose. In *1-out-of-2 chosen k-bit-string OT* ($\binom{2}{1}$ k-bit-string OT) the messages are k-bit strings instead of single bits. One reason for the importance of OT is that it allows for carrying out *any* secret two-party computation.

Many results, lower as well as upper bounds, have been proven on information-theoretic reductions of such OT primitives to others [5], [4], [6], [7]. It is a particularly interesting problem to realize OT (e.g., chosen string OT) from primitives that are as weak as possible [6]. It is the purpose of this paper to perform another step into this direction.

2 Bit-, String-, and Generalized OT: Definitions and Previous Results

Many approaches have been taken for reducing *string* OT to *bit* OT. Such a reduction can be based on so-called *zig-zag functions* or on *self-intersecting codes* (see [3] and the references therein). The drawback of these approaches is that the number of required bit-OT realizations cannot be made smaller than about $3.538k$, where k is the length of the strings. The *privacy-amplification* method, presented in [3], achieves roughly $2k$, but the resulting string OT must be tolerated to fail with a certain probability.

Privacy amplification was first considered in the context of information-theoretic secret-key agreement [2] and is the technique of transforming a partially secret into a highly secret string. Equivalently, privacy amplification can be seen as "distribution uniformizing." A good technique for privacy amplification was shown to be universal hashing. In this case, the length of the highly secret key that can be extracted is roughly equal to the Rényi entropy of the original string [1].

The reduction of string- to bit OT by privacy amplification can be done by the following protocol of [3]. Note that privacy amplification means here applying a random linear function. The protocol reduces $\binom{2}{1}$ k-bit-string OT of the k-bit messages m_0 and m_1 to n realizations of $\binom{2}{1}$ OT with some failure probability. A discussion of the parameters is given below. Let $\mathrm{lin}(2^n, 2^r)$ denote the set of linear functions mapping n-bit strings to r-bit strings, and let $\in_r$ stand for the random choice of an element out of a set according to the uniform distribution. Let furthermore a^i denote the i-th coordinate of the vector a.

Protocol BC97

Alice		**Bob**
$x_0, x_1 \in_r GF(2^n)$		$c \in \{0,1\}$
	$\binom{2}{1}$ OT (x_0^i, x_1^i) $\longrightarrow$	
		$x_c^1, x_c^2, \ldots$
$h_0, h_1 \in_r \mathrm{lin}(2^n, 2^r)$		
$y_0 := m_0 \oplus h_0(x_0)$		
$y_1 := m_1 \oplus h_1(x_1)$		
	h_0, h_1, y_0, y_1 $\longrightarrow$	
		$m_c = h_c(x_c) \oplus y_c$

It was shown in [3] that for $n \geq 2k+s$, Protocol BC97 works with a failure probability of 2^{-s} (more precisely, there exists after the protocol execution, with probability at least $1 - 2^{-s}$, a bit c such that Bob has no information at all about the message m_{1-c}, even when given m_c).

More generally, the same was even shown to be true when $\binom{2}{1}$ OT is replaced by so-called $\binom{2}{1}$ XOT (b_0, b_1), where Bob can not only choose to obtain the bits b_0 or b_1, but also $b_\oplus := b_0 \oplus b_1$. The primitive XOT was further extended to *generalized OT* (GOT) allowing, besides the unbiased functions

$$\begin{aligned} 0(b_0, b_1) &:= b_0 \ , \\ 1(b_0, b_1) &:= b_1 \ , \\ \oplus(b_0, b_1) &:= b_\oplus \end{aligned}$$

(and their negations) also the *biased* binary functions $\vee, \wedge, \rightarrow, \leftarrow$ (and their negations) of two bits as possible choices of Bob. The price that has to be paid for this generalization is a reduced efficiency of the reduction. Let us first give a definition of the security of string OT with failure probability.

Definition 1 A $\binom{2}{1}$ *k-bit-string OT with security s* has the property that there exists an event $\mathcal{A}$ with probability at least $1 - 2^{-s}$, taken over all possible choices of Bob and over all the coin tosses of Alice, such that given that $\mathcal{A}$ occurs, the receiver Bob obtains no information about one of the k-bit strings, even when given the other.

Remark. Note that string OT with security s means here that Bob obtains no information at all about at least one of the strings with high probability. In [3], the corresponding definition was somewhat weaker and required that with probability at least $1 - 2^{-s}$, Bob will find himself in a situation where he knows at most 2^{-s} Shannon bits about one of the two strings, given the other. By the side-information argumentation needed in the proof of Lemma 2 below one can generally show that the two definitions are equivalent in principle.

The reduction of $\binom{2}{1}$ k-bit-string OT with security s to the new primitive of GOT was shown in [3] to require $O(k + s)$ executions of GOT.

3 The Power of Universal OT

The most general (i.e., weakest) primitive in the described context appears to be the so-called *universal* OT proposed in [3]. Here, Bob is allowed to choose *any* type of information, in particular probabilistic information, about the bits sent by Alice, not exceeding a certain bound on Shannon entropy. Obviously, this primitive is much more general than GOT. For instance, Bob

can choose here to receive slightly noisy versions of both bits b_0 and b_1 (with some arbitrarily small error probability ε).

Definition 2 Let $\alpha < 2$. A *universal oblivious transfer with parameter* α (α-UOT for short) is a cryptographic primitive involving two parties Alice (called *sender*) and Bob (the *receiver*). The sender Alice's input is a pair of bits (b_0, b_1). The receiver Bob on the other hand inputs a (possibly probabilistic) function Ω which must satisfy

$$H((B_0, B_1) \,|\, \Omega(B_0, B_1)) \geq \beta$$

(where $\beta := 2 - \alpha$). Furthermore, the receiver obtains $\Omega(B_0, B_1)$, but no additional information about (B_0, B_1). Finally, Alice does not learn anything about Bob's choice of the function Ω.

It was stated as an open problem in [3] whether this primitive is as strong as string OT, i.e., whether it is also possible to efficiently reduce $\binom{2}{1}$ k-bit-string OT to general UOT. Theorem 1 shows that the answer to this question is *yes*, and that the number of required realizations of α-UOT (for any fixed $\alpha < 2$) is of order $O(k+s)$. The proof of Theorem 1 given below is somewhat simpler than the analysis of the GOT reduction in [3].

Theorem 1 *Protocol BC97 reduces* $\binom{2}{1}$ *k-bit-string OT with security s to n realizations of α-UOT for every*

$$n \geq \left\lceil \frac{(s + 2k) \cdot 2 \ln 2}{p_e} \right\rceil ,$$

where p_e is the unique solution ($\leq 1/2$) to the equation

$$h(x) + x \log 3 = 2 - \alpha =: \beta .$$

The statement of Theorem 1 is meant in the sense that it also holds if the receiver chooses the new function Ω adaptively after each step.

Before proving Theorem 1, we need the following lemma which intuitively implies that among all possible types of partial information about a bit with a fixed error probability about this bit, the particular information that is obtained by sending the bit over a symmetric erasure channel provides the largest amount of Shannon information about the bit. Even stronger than that, we show that for every other type of information, there exists side information V (that can be thought of as being provided by an oracle) such that given this information in addition, the situation perfectly

corresponds to information obtained from an erasure channel. Note that this side-information argumentation leads to a partial order on all possible types of side information about a random variable in a very strict sense: the "stronger" side information is "more powerful" than the weaker one in every respect because the stronger information contains the weaker one.

Lemma 2 *Let B be a symmetric binary random variable, and let U be a random variable such that B and U have joint distribution P_{BU}. Let p be the average error probability of guessing B when given U, using the optimal guessing strategy. Then there exists a random variable V with the following properties:*

1. $\mathcal{V} = \{0, 1, \Delta\}$,
2. $P_V(\Delta) = 2p$,
3. *for every $u \in \mathcal{U}$, we have*

$$P_{B|U=u,V=\Delta}(0) = P_{B|U=u,V=\Delta}(1) \ .$$

Proof. Let $u \in \mathcal{U}$, and assume that $a = P_{B|U=u}(0) > P_{B|U=u}(1) = b$. Let V be defined by

$$\begin{aligned} P_{V|B=0,U=u}(0) &= (a-b)/a \ , \\ P_{V|B=0,U=u}(\Delta) &= b/a \ , \\ P_{V|B=1,U=u}(\Delta) &= 1 \ . \end{aligned}$$

Note that $P_{V|U=u}(\Delta) = 2b$, i.e., twice the error probability for guessing B when given $U = u$. This concludes the proof. □

Proof of Theorem 1. Let n be the length of the strings x_0 and x_1 in Protocol BC97. According to Fano's inequality, the expected error probability, given $\Omega_i(x_0^i, x_1^i)$, about the pair of bits (x_0^i, x_1^i) is at least p_e, where p_e stands for the unique solution ($\leq 1/2$) to the equation $h(x) + x \cdot \log 3 = \beta$ (h is the binary entropy function). This means that about at least two of the bits x_0^i, x_1^i, $x_\oplus^i$ ($:= x_0^i \oplus x_1^i$), the expected error probability is at least $p_e/2$.

Let $g(\cdot, \cdot)$ be a linear function mapping $[GF(2)^r]^2$ to $GF(2)$ depending non-trivially on both inputs (i.e., it is not true that either $g(x, y) = g(x', y)$ or $g(x, y) = g(x, y')$ holds for all $x, x', y, y' \in GF(2)^r$). We consider the probability that Bob can bias the bit

$$g(h_0(x_0), h_1(x_1)) \ . \tag{1}$$

For every $i = 1, \ldots, n$, the bit (1) can be written as

$$a_0 x_0^i \oplus a_1 x_1^i \oplus R_i(x_0^1, \ldots, x_0^{i-1}, x_0^{i+1}, \ldots, x_0^n, x_1^1, \ldots, x_1^{i-1}, x_1^{i+1}, \ldots, x_1^n) \ ,$$

where $a_0, a_1 \in \{0,1\}$ are independent and random (given that g depends non-trivially on both input strings and that h_0 and h_1 are independent and random), and where R_i is a linear function mapping to a bit.

We conclude from the above that with probability at least $1-(1/4+3/4\cdot 1/3) = 1/2$, Bob's expected error probability about the bit $a_0 x_0^i \oplus a_1 x_1^i$ he needs is at least $p_e/2$, hence his total expected error probability is at least $p_e/4$. As Lemma 2 shows, the worst case (for Alice) is when Bob has full information about the required bit with conditional probability $1-p_e/2$, and no information otherwise. Thus Bob will in this case have *no information at all* about $g(h_0(x_0), h_1(x_1))$ with probability

$$1 - (1 - p_e/2)^n \ .$$

Hence the probability $\mathrm{Prob}\,[\mathcal{E}]$ of the event $\mathcal{E}$ that there exists a non-trivial bilinear function g such that $g(h_0(x_0), h_1(x_1))$ can be non-trivially biased by Bob is, by the union bound, bounded by

$$\mathrm{Prob}\,[\mathcal{E}] < 2^{2k}(1 - p_e/2)^n \ .$$

It is not difficult to see that, given the event $\mathcal{A} := \overline{\mathcal{E}}$ that $\mathcal{E}$ does *not* occur, we have that S_0 and S_1 are statistically independent, and that at least one of them is perfectly uniformly distributed (see [3] for a result implying that if every non-trivial linear function of a string is unbiased then the string is perfectly uniform).

Therefore, about one of the resulting strings, Bob has no information at all with probability at least

$$1 - 2^{2k}(1 - p_e/2)^n \ , \tag{2}$$

given the other string. Expression (2) is at least $1 - 2^{-s}$ for all

$$n \geq \left\lceil \frac{(s + 2k) \cdot 2 \ln 2}{p_e} \right\rceil \ ,$$

and this concludes the proof. □

References

[1] C. H. Bennett, G. Brassard, C. Crépeau, and U. M. Maurer, Generalized privacy amplification, *IEEE Transactions on Information Theory*, Vol. 41, No. 6, pp. 1915–1923, 1995.

[2] C. H. Bennett, G. Brassard, and J.-M. Robert, Privacy amplification by public discussion, *SIAM Journal on Computing*, Vol. 17, pp. 210–229, 1988.

[3] G. Brassard and C. Crépeau, Oblivious transfers and privacy amplification, *Advances in Cryptography – EUROCRYPT' 97*, LNCS, Vol. 1233, pp. 334–345, Springer-Verlag, 1997.

[4] C. Cachin, On the foundations of oblivious transfer, *Advances in Cryptography – EUROCRYPT' 98*, LNCS, Vol. 1403, pp. 361–374, Springer-Verlag, 1998.

[5] C. Crépeau, *Correct and private reductions among oblivious transfers*, PhD thesis, MIT, 1990.

[6] I. Damgård, J. Kilian, and L. Salvail, On the (im)possibility of basing oblivious transfer on weakened security assumptions, *Advances in Cryptography – EUROCRYPT' 99*, LNCS, Vol. 1592, pp. 56–73, Springer-Verlag, 1999.

[7] Y. Dodis and S. Micali, Implementing oblivious transfer, *Advances in Cryptography – EUROCRYPT' 99*, LNCS, Vol. 1592, pp. 42–55, Springer-Verlag, 1999.

[8] U. Maurer, Information-theoretic cryptography, *Advances in Cryptography – CRYPTO' 99*, LNCS, Vol. 1666, pp. 47–64, Springer-Verlag, 1999.