# Secure Stochastic Multi-party Computation for Combinatorial Problems and A Privacy Concept that Explicitely Factors out Knowledge about the Protocol ${ }^{*}$ 

Marius C. Silaghi ${ }^{\dagger}$ and Gerhard Friedrich ${ }^{\ddagger}$<br>${ }^{\dagger}$ Florida Institute of Technology, USA<br>${ }^{\ddagger}$ University Klagenfurt, Austria

January 2, 2006


#### Abstract

High levels of security often imply that the computation time should be independent of the value of involved secrets. When the expected answer of the solver is either a solution or unsatisfiable, then the previous assumption leads to algorithms that take always the computation time of the worst case. This is particularly disturbing for NP-hard combinatorial problems.

In this work we start from the observation that sometimes (specially for hard problems) users find it acceptable to receive as answer either a solution, the answer unsatisfiable or a failure with meaning don't know. More exactly users accept incomplete solvers. As argued in [Sil05b], for certain problems privacy reasons lead users to prefer having an answer meaning don't know even when the secure multi-party computation could have proven unsatisfiable (to avoid revealing that all alternatives are infeasible). While the solution proposed in [Sil05b] is slower than complete algorithms, here we show secure stochastic solutions that are faster than complete solvers, allowing to address larger problem instances. Two new refined concepts of privacy are introduced, namely requested t-privacy that factors out treatment of knowledge of the protocol in t-privacy, and a slightly weaker version called non-uniform requested t-privacy. In the last section we discuss arithmetic circuits for complete and stochastic solutions to constraint optimization problems.


## 1 Introduction

Typical examples of combinatorial problems are meeting scheduling, resource allocation, time-tabling, auctions with several possible winners. Such a problem is typically

[^0]defined by a set of variables and constraints on the satisfiable assignments to these variables. The set of all (satisfiable and unsatisfiable) simultaneous assignments of values to all variables defines the search space of the problem. An element of the search space is also referred to as an alternative to be considered as a solution to the problem, or simply alternative.

A complete solver is one that reports a solution whenever a solution exists. The answer of such a technique is either a solution or unsatisfiable. Combinatorial problems can be very hard and therefore we no not have efficient complete secure multi-party computation solvers. Several complete secure solvers were proposed in the past for such problems, and high levels of security always require a computation time that is given by the worst possible case (over all possible values of the secrets).

It was shown that for problems that are solved only once, minimization of privacy loss often requires that the solution be picked randomly, preferably with a uniform distribution among the existing solutions [SR04]. Such a random selection can be achieved if the problem is shuffled prior to solving [Sil03, Sil04]. Two families of techniques were proposed for shuffling a shared description of a combinatorial problem,one based on mix-nets and one based on arithmetic circuits [Sil05c].

Sometimes, the security requirements themselves require an incomplete solver (when the proof of unsatisfiability of the problem leads to unacceptable privacy loss, by revealing that all alternatives are infeasible) [Sil05b]. The answer of such a solver is either a solution or unsatisfiable. However, the solution proposed in [Sil05b] is actually slower than complete solutions. It first computes a solution with a complete secure solver and then it hides the solution with some small probability.

In this work we show how the shuffling performed on problem descriptions prior to solving allows to build an incomplete secure stochastic multi-party solver where a high level of privacy is offered. The answers of the solver consists in either a solution or in don't know, and nothing is revealed about the set of alternatives that were not explored (except for its size). Notably, these algorithms are strictly faster than the corresponding complete versions and are parametrized with the percentage of the search space to be explored (the search space is the set of all alternatives that may or may not satisfy the combinatorial problem).

By specifying the percentage of the combinatorial problem to be explored, one practically specifies the exact amount of computation (time) that the solver should perform. The proposed techniques are different for shuffling with mix-nets and for shuffling with arithmetic circuits. Arithmetic circuits for optimization problems are discussed in the last sections.

## 2 Privacy Concept that Explicitely Factors out Knowledge about the Protocol

We refine a concept of privacy that allows to formally identify losses of privacy due to certain multi-party computation protocols. Namely, previous privacy concepts suggest that a protocol is secure if everything leaked by its execution can also be inferred from prior knowledge and obtained result. However, in this article we stress the need to
classify the information leaked due to the knowledge of which protocol is employed as being leaked by its execution rather than solely by prior knowledge and obtained result. This distinction is non-trivial, as proven by the fact that several recent articles present algorithms designed according to the aforementioned privacy concept and still failing our test of security.

The implications of the two ways of accounting privacy loss becomes easily clear for problems where several solutions are possible (optimal). For such problems, the knowledge of the algorithm used to select the solution may leak information about where the other (optimal) solutions may or may not be located. While these implications were underlined in our previous work, the new concept of privacy shown here offers a formal way to evaluate solutions proposed in the past for this problem. This allows to motivate protocols leaking less information for applications such as auctions.

Different privacy concepts and levels of privacy were identified for multi-party computations [BOGW88, FMW01], and they are relevant for different computational models and assumptions. The t-privacy concept introduced in [BOGW88] is very well known and is applicable to many types of assumptions:

Definition 1 A multi-party computation is t-private if an attacker controlling any at most $t$ participants cannot learn anything from the computation, except for what can be inferred from its outputs and prior knowledge.

Security of a t-private scheme can be computational (if breaking it is computationally intractable for the attacker) or information theoretical (if even infinite computation power cannot infer anything about secrets since no information about it is contained in the data obtained by the attacker).

Given secret inputs $\sigma$, the prior knowledge $\Gamma$ of the $t$ colluders and a multi-party computation process $\Pi$ with answer $\alpha$, the technique is t-private if the probability distribution of the secrets is conditionally independent on $\Pi$ given answer $\alpha$ and knowledge $\Gamma$.

$$
\begin{equation*}
P(\sigma \mid \alpha, \Gamma, \Pi)=P(\sigma \mid \alpha, \Gamma) \tag{1}
\end{equation*}
$$

However, many algorithms provide answers $\alpha$ that contain more information than what is actually needed. We typically decompose $\alpha$ in a desired data $\alpha^{*}$ and an algorithmic dependent unrequested data $\bar{\alpha}$. For example ${ }^{1}$ the desired data can be an assignment of some variables satisfying secret constraints, and the unrequested data is what can be obtained from peculiarities of the used algorithm $\mathcal{A}$ (e.g., the solution is the first/last in some known order on alternatives).

Definition 2 We say that an algorithm $\mathcal{A}$ achieves requested t-privacy if the probability distribution of the secrets is conditionally independent on $\Pi, \mathcal{A}$ and $\bar{\alpha}$ given requested data $\alpha^{*}$ and prior knowledge $\Gamma$.

$$
\begin{equation*}
P(\sigma \mid \alpha, \Gamma, \Pi, \mathcal{A})=P\left(\sigma \mid \alpha^{*}, \Gamma\right) \tag{2}
\end{equation*}
$$

[^1]Examples of algorithms designed to be secure, according to Equation 1 [BOGW88], but insecure according to the definition introduced in Equation 2 are proposed in [YSH02b, Sil03, YS04, NZ05].

For problems with several solutions, requested t-privacy typically implies the return of uniformly random selected solutions whenever the problem may have more than one solution. Examples of techniques secure according to the introduced concept are [Sil04, Sil05a]. Sometimes uniform randomness in selecting the solution requires very expensive computations. Non-uniform randomness in selecting the solution, while shown to be often clearly better then deterministic approaches and often achievable with significantly reduced computation effort, is less secure [Sil04]. The level of privacy achieved in that case is nevertheless interesting and worth its own definition. We propose to call it non-uniform requested t-privacy, and can be describes by the property:

Definition 3 We say that an algorithm $\mathcal{A}$ achieves non-uniform requested t-privacy if for any secret $\bar{\sigma} \in \sigma$ that is not deterministically revealed given requested data $\alpha^{*}$ and prior knowledge $\Gamma$, it is also not deterministically relealed given $\Pi, \mathcal{A}$ and $\bar{\alpha}$.

$$
\begin{equation*}
\forall \bar{\sigma} \in \sigma P\left(\bar{\sigma} \mid \alpha^{*}, \Gamma\right)<1 \Rightarrow P(\bar{\sigma} \mid \alpha, \Gamma, \Pi, \mathcal{A})<1 \tag{3}
\end{equation*}
$$

## 3 Background

Combinatorial problems have been often discussed in Computer Science and many examples are known to be very hard. For example SAT was the first proven NP-complete problem and Constraint Satisfaction Problems are largely addressed with stochastic and incomplete solvers.

A Constraint Satisfaction Problem (X,D,C) is defined by a set of variables $X=$ $\left\{x_{1}, \ldots, x_{m}\right\}$, a set of domains $D=\left\{D_{1}, \ldots, D_{m}\right\}$ where $D_{i}$ is the domain for $x_{i}$, and a set of constraints $C=\left\{\phi_{1}, \ldots, \phi_{c}\right\}$. Each constraint $\phi_{j}$ specifies the acceptable combinations of assignments of values to a subset $X_{j}$ of the variables. A tuple is a vector of assignments of values to distinct variables. A solution of the CSP is a tuple of assignments of values to all the variables and that satisfies all the constraints. The search space of the CSP is defined by the Cartesian product $D_{1} \times \ldots \times D_{m}$. An element of the search space is called an alternative. The $i^{t h}$ alternative is denoted by $\epsilon_{i}$.

A distributed CSP is a $\operatorname{CSP}(X, D, C)$ where a set of participants $A=\left\{A_{1}, \ldots, A_{n}\right\}$ have secret shares of $C$, none of them knowing the whole set $C$.

### 3.1 Shuffling an array of shared secrets

Secure multi-party computations can simulate any arithmetic circuit [BOGW88] or boolean circuit [Kil88, Gol04] evaluation. An arithmetic circuit can be intuitively imagined as a directed graph without cycles where each node is described either by an addition/subtraction or by a multiplication operator. Each leaf is a constant.

The secure multi-party simulation of arithmetic circuit evaluation proposed in [BOGW88] exploits Shamir's secret sharing [Sha79]. This sharing is based on the
fact that a polynomial $f(x)$ of degree $t-1$ with unknown parameters can be reconstructed given the evaluation of $f$ in at least $t$ distinct values of $x$, using Lagrange interpolation. Absolutely no information is given about the value of $f(0)$ by revealing the valuation of $f$ in any at most $t-1$ non-zero values of $x$. Therefore, in order to share a secret number $s$ to $n$ participants $A_{1}, \ldots, A_{n}$, one first selects $t-1$ random numbers $a_{1}, \ldots, a_{t-1}$ that will define the polynomial $f(x)=s+\sum_{i=1}^{t-1}\left(a_{i} x^{i}\right)$. A distinct non-zero number $\tau_{i}$ is assigned to each participant $A_{i}$. The value of the pair $\left(\tau_{i}, f\left(\tau_{i}\right)\right)$ is sent over a secure channel (e.g. encrypted) to each participant $A_{i}$. This is called a $(t, n)$-threshold scheme. We will assume that all computations are performed in a field $\mathbb{Z}_{q}$ for some prime number $q$. Once secret numbers are shared with a $(t, n)$-threshold scheme, evaluation of an arbitrary arithmetic circuit can be performed over the shared secrets, in such a way that all results remain shared secrets with the same security properties (the number of supported colluders, $t-1$ ) [BOGW88, Yao82]. For [Sha79]'s technique, one knows to perform additions and multiplications when $t \leq(n-1) / 2$. Since any $\lfloor n / 2\rfloor$ participants cannot find anything secret by colluding, such a technique is called $\lfloor n / 2\rfloor$-private [BOGW88]. It is also known how to evaluate with computational securely any arithmetic circuit on additively shared secrets.

Shuffling with mix-nets In [Sil03, Sil04, Sil05c] it is shown how a mix-net can shuffle a vector of shared secrets and can unshuffle a vector of the same size using the inverse permutations. Each participant encrypts his share of each secret using a $(+\bmod q, X)$ public encryption scheme for which it holds the secret key, and sends a vector holding each encrypted share to $A_{1}$. The vectors with the encrypted shares are passed along each participant in $A$, each of the applying the same secret permutation on all vectors. A shared 0 is also added to each sharing of a secret using the homomorphism of the encryption. Each participant will provide the others with a zeroknowledge proof for the correctness of his shuffling (respectively unshuffling).

Shuffling with arithmetic circuits Assume that we have composable multi-party computations [Kil05, DFNT05] for computing:

- $\delta_{K}(x, y)$ : Kronecker's delta returning a shared 1 when $x=y$ and 0 otherwise
- $\operatorname{cmp}(x, y)$ returns 1 when $x<y$ and 0 otherwise
- $R S(m, M)$ : random secret generator, generating a shared secret in the interval $m, M$.

It is possible to design an arithmetic circuit for shuffling secrets, using the Algorithm 3. This algorithm uses Algorithm 1 for a permutation of two elements on secret positions in a vector. The random permutation is defined by a random vector computed with Algorithm 2. Unshuffling can be done with the Algorithm 4.

This permutation was shown in [Sil05c] to lead to a random shuffling (taken from a uniform distribution). Note that the random vector defining the permutation could have been built allowing each element to belong to any value between 1 and $k$. This would be computationally more expensive as it would require each call to the procedure Perm to recompute all the elements of the vector to be shuffled (see Algorithm 5).

```
function \(\operatorname{Perm}(s, i, r, m, M, k\)
    \(s_{i}=\sum_{j=m}^{M}\left(\delta_{K}(r, j) * s_{j}\right) ;\)
    for \(j \in(i, k]\) do
        \(s_{j}=s_{j}+\left(s_{i}-s_{j}\right) * \delta_{K}(r, j) ;\)
```

Algorithm 1: Permuting element $s_{i}$ with $s_{r}$ for a secret value $r \in[m, M]$ in vector $s$ with $k$ shared secrets

```
function RandomVector(k)
        for j=1 to k-1 do
        r[j] = RS(j,k);
    return r;
```

Algorithm 2: Shuffling a vector $s$ with $k$ shared secrets

```
function Shuffle(s,k,r)
    for }j=1\mathrm{ to }k-1\mathrm{ do
        Perm(s,j,r[j],j,k,k);
```

Algorithm 3: Shuffling a vector $s$ with $k$ shared secrets, and a random vector $r$ obtained with Algorithm 2

```
function Shuffle(s,k,r)
    for }j=k-1 to 1 do
        Perm(s,j,r[j],j,k,k);
```

Algorithm 4: Un-shuffling a vector $s$ with $k$ shared secrets, when the shuffling was defined by random secret vector $r$.

```
function Shuffle(s,k,r)
    for }j=1\mathrm{ to }k\mathrm{ do
        Perm(s,j,r[j],1,k,k);
```

Algorithm 5: Shuffling a vector $s$ with $k$ shared secrets, and a random vector $r$ where each element is obtained with $R S(1, k)$.

### 3.2 MPC-DisCSP4

In [Sil05b] we have proposed a multi-party computation technique, called MPCDisCSP4, that extracts a random solution of a distributed CSP. MPC-DisCSP4 uses general multi-party computation building blocks. General multi-party computation techniques can solve securely certain functions, one of the most general classes of solved problems being the arithmetic circuits. A distributed CSP is not a function. A DisCSP can have several solutions for an input problem, or can even have no solution. Two of the three reformulations of DisCSPs as a function (see [SR04]) are relevant for MPC-DisCSP4:
$i$ A function DisCSP ${ }^{1}()$ returning the first solution in lexicographic order, respectively an invalid valuation $\tau$ when there is no solution.
ii A probabilistic function DisCSP() which picks randomly a solution if it exists, respectively returns $\tau$ when there is no solution.

For privacy purposes only the $2^{\text {nd }}$ alternative is satisfactory. $\operatorname{DisCSP}()$ only reveals what we usually expect to get from a DisCSP, namely some solution. DisCSP ${ }^{1}$ () intrinsically reveals more [SR04]. MPC-DisCSP4 implements DisCSP() in five phases:

1. Share the secret parameters of the input DisCSP using Shamir's secret sharing. The value of each publicly possible assignment (allocation) is securely evaluated.
2. The shared DisCSP problem is shuffled in a cooperative way, reordering values (and eventually variables), with a permutation that is not known to anybody [Sil05c].
3. A version of $\operatorname{DisCSP}^{1}()$ where the operations performed by agents are independent of the input secrets (to avoid leaking the secrets), is executed by simulating arithmetic circuits evaluation with the technique in [BOGW88].
4. The solution returned by $\operatorname{DisCSP}^{1}()$ at Step 3 is translated into the initial problem formulation using a transformation that is inverse of the shuffling at Step 2 [Sil05c].
5. Construct the solution from its secret shares.

It is also possible and very simple to find all solutions $\left[\mathrm{HCN}^{+} 01\right]$. However, when only a single solution is needed, this leaks a lot of information. At Step 3, MPCDisCSP4 requires a version of the DisCSP ${ }^{1}$ () function whose cost is independent of the input, since otherwise the users can learn things like: The returned solution is the only one, being found after unsuccessfully checking all other tuples, all other tuples being infeasible. Since the used DisCSP ${ }^{1}$ () has to be independent of the problem details, its cost is exponential (at least as long as nobody proves $\mathrm{P}=\mathrm{NP}$ ).

Note that other alternative techniques are available, notably MPC-DisCSP1 [Sil03], MPC-DisCSP2 [SM04], and MPC-DisCSP3 [Sil04]. We call them generically MPCDisCSPx. In this paper we only address multi-party computations without trusted servers. A family of secure solvers based on trusted servers is proposed in [YSH02a].


Figure 1: MPC-DisCSP4 using mix-nets

### 3.3 Hiding existence of solution

When no solution is found, all the participants learn that each alternative is infeasible. For certain problems this leak of secrets may be considered unacceptable and a don't know answer is prefered to learning the infeasibility. But the don't know answer is believable only if the algorithm may indeed miss some solutions. An algorithm for missing the solution with some predefined probability $p$ is described in [Sil05b]. It consists of computing a solution using a MPC-DisCSPx algorithm and then setting the assignments in the result to the invalid value 0 with a probability $p$.

### 3.4 Stochastic algorithms

In the CSP world it is known that complete algorithms are ineffective for hard problem instances. For large problems, most applications apply stochastic search procedures. With stochastic search, only a subset of the search space is analyzed. Typical examples of stochastic search are based on some type of hill climbing. With hill-climbing the solver starts with a random alternative and searches the neighbouring search space for solutions.

## 4 Simulated Annealing for Secure Optimization

Once the secret constraints of a distributed CSP are shared and shuffled with the technique of MPC-DisCSP1 [Sil03, Sil05c], one can try to search a feasible solution of the shuffled problem using some hill-climbing. The same considerations and procedures apply if the problem is shuffled with a mixnet obtained from the one in [YSH02a] by replacing the encryption scheme with a $(+, \times)$-homomorphic version (E.g., Paillier with shared secret key, or the version of ElGamal of the form $E_{a, y, g, p}(m, r)=\left\langle g^{r}\right.$ $\left.\bmod p, a^{m} y^{r} \bmod p\right\rangle$ ).

The quality of an alternative will normally be evaluated securely (since we do not tipically want to reveal individual constraints even if they were shuffled - as it would lead to an important privacy loss). The total weight (or number of conflicting constraints) for an alternative $\epsilon$ is computed with $q(\epsilon, P)=\sum_{c \in C} c(\epsilon)$. The revelation of the quality will be relatively expensive for both versions (based on either secret sharing or homomorphic encryption). Therefore, this suggests to use stochastic algorithms that are lazy in evaluating the qualities of new tuples. Such a technique is Simulated Annealing (Algorithm 6).

```
procedure SSA do
    Shuffle DisCSP using secret sharing or additive encryption homomorphism;
    Select random alternative (tuple) t;
    for decreasing 'temperature' }T\mathrm{ do
        change randoly the value of one variable obtaining t';
        compute securely and then reveal }\Delta=q(\mp@subsup{t}{}{\prime})-q(t)
        /*alternatively reveal q(t') to detect termination when the optimum is
        known*/;
        /*or securely compute and reveal only }\operatorname{cmp}(q(\mp@subsup{t}{}{\prime}),q(t)), if it returns 1*/
        if }\Delta<0\mathrm{ then
            t=t'
        else
            t=t'}\mathrm{ with probability }\mp@subsup{e}{}{-\frac{\Delta}{T}
    Unshuffle the results;
```

                Algorithm 6: Secure Simulated Annealing (minimization)
    Similar to the technique in [Sil02, YSH02a], the Secure Simulated Annealing algorithm may reveal undesired statistical information about some secrets via the knowledge of the shuffle search space. However, specific exact information about a secret may only be inadvertently revealed only for problems with very special patterns. In the following we concentrate on algorithms guaranteed not to reveal anything else besides the solution.

## 5 Secure Stochastic Search

Let us finally detail our proposed techniques for tractable secure stochastic search, allowing to address hard problems. The idea is that only a subset of $T$ alternatives from, the search space will be explored. This could be achieved by adding a public constraint that removes the remaining search space. However, to ensure privacy in case of failure (that the infeasibility of this particular sub-space is not revealed), we propose to take advantage of the shuffling of the whole problem. We select the subspace to be explored from the shuffled problem. This hides the exact search subspace that is analyzed and the only secret leaked in case of failure is that there are $T$ infeasible alternatives (but they are not known).

### 5.1 Secure Stochastic Search with Mix-nets

Each MPC-DisCSPx solving algorithm using mixnets can be modified into a corresponding secure stochastic search protocol that will be called Stochastic Multi-Party Computation for Distributed CSPs (SMPC-DisCSPx). Each SMPC-DisCSPx differs from the corresponding MPC-DisCSPx by the fact that only the first $T$ tuples of the shuffled search space are used to compute the shuffled solution. Each stochastic solver is parametrized by the number $T$ of alternatives to be explored ( $T$ beeing smaller or equal to the size of the search space). To be noted that a stochastic solver can be seen as a generalization of the corresponding complete solver, which is obtained when $T$ equals the size of the search space.

SMPC-DisCSP4 For example, SMPC-DisCSP4 is shown in Algorithm 7.

```
    function SMPC-DisCSP4(T,(X,D,C))
        for }i=1\mathrm{ to }k\mathrm{ do
        S[i]=\}\mp@subsup{\prod}{\phi\inC}{}\phi(\mp@subsup{\epsilon}{i}{})
    SHUFFLE(S) //using the mixnet;
    h[1]=1;
    for i=2 to T do
        h[i]=h[i-1]*(1-S[i-1]);
        S[i]=S[i]*h[i];
    /* S[T]=S[T]*cmp(RS(0,q-1),p*q)// fine tuning*/;
    UNSHUFFLE(S);
    set solution S to 0 with probability p; //optional;
    return S// the solution can be extracted from S as in [Sil05b];
```

Algorithm 7: SMPC-DisCSP4 for solving a CSP $(X, D, C)$ with $k$ alternatives allowed by the public constraints, and exploring $T$ alternatives.

SMPC-DisCSP4 requires $k(c-1)$ multiplications of secrets to build the vector $S$ and $2 T$ multiplications of secrets to select the solution. Also, the shuffling and unshuffling require each $O\left(k n^{2}\right)$ expensive operations, $O(k n)$ for each participant. While SMPC-DisCSP4 leads to a reduction with up to $2 k$ multiplications of secrets, the complexity remains the same, dictated by the shuffling.

It can be noted that the probability that a solution is lost can be fine tuned (e.g. for the application in [Sil05b]) by discarding the alternative $\epsilon_{T}$ with probability $p$. This can be done by uncommenting the Line 7.1 in the Algorithm 7.

One can allow agents to avoid revealing if that there exist $T$ alternatives that are not solutions, by enabling the optional cancelation of the solution with probability $p$ at Line 7.1. This cancelation of solution can be done with the technique in [Sil05b]).

SMPC-DisCSP1 The stochastic algorithm obtained from MPC-DisCSP1 is more successful, and is sketched in Algorithm 8.

DisCSP1 (Figure 2) is the arithmetic circuit proposed in [Sil03], with the only modification that function satisfiable () only integrates the first $T$ tuples (rather

```
function SMPC-DisCSP1(T,(X,D,C))
    SHUFFLE(X,D,C) //using the mixnet;
    for }i=1 to T d
        S[i]= \}\mp@subsup{|}{\phi\inC}{}\phi(\mp@subsup{\epsilon}{i}{})
    F=DisCSP1(T,(X,D,C));
    UNSHUFFLE(F); // Unshuffle each vector in F separately;
    set solution F to 0 with probability p; //optional;
    return F;
```

Algorithm 8: SMPC-DisCSP1 for solving a CSP $(X, D, C)$ with $k$ alternatives and exploring $T$ alternatives.

$$
\begin{aligned}
p(\epsilon, P) & =\prod_{c \in C} c(\epsilon) \\
\text { satisfiable }(P) & =c m p\left(0, \sum_{\epsilon_{i} \in\left[\epsilon_{1} \ldots \epsilon_{T}\right]} p\left(\epsilon_{i}, P\right)\right) \\
g_{i, j}(P) & =\operatorname{satisfiable}\left(P \cup\left\{x_{i}=j\right\} \cup_{k<i}\left(x_{k}=f_{k}(P)\right)\right) \\
f_{j}(P) & =\sum_{i=1}^{\left|D_{j}\right|} i *\left(g_{j, i}(P) * \delta_{K}\left(0, \sum_{k<i} g_{j, k}(P)\right)\right)
\end{aligned}
$$

Figure 2: Arithmetic circuit DisCSP1 for a CSP $P=(X, D, C)$. The result is the vector of vectors $\left\{\left\{\delta_{K}\left(f_{i}, j\right)\right\}_{j \in\left[1 . .\left|D_{i}\right|\right]}\right\}_{i \in[1 . . m]}$. Versions with other primitives appear in [Sil03, Sil04]
than the whole search space). The result $F$ returned by DisCSP1 is a set of vectors, one for each variable. A vector contains shared 0 s on all positions, except for a 1 on the position corresponding the the value of the corresponding variable in the found solution. If there is no solution, then all elements of the vectors are 0 .

The cost of SMPC-DisCSP1 is only $O(T(m d+c))$ multiplications of secrets. Of these, $T(c-1)$ are used to compute $S$. DisCSP1 computes satisfiable $m d$ times, each of them requiring at most $O(T)$ multiplications. The cost of shuffling in SMPCDisCSP1 can be small even for large and hard problems, if the maximum constraint arity (number of involved variables) is small.

The stochastic techniques can be adapted to the corresponding secure optimization techniques based on secure distributed weighted CSPs, MPC-DisWCSPx [SM04], leading to algorithms that we will denote SMPC-DisWCSPx. When using such optimization techniques for applications like Generalized Vickerey Auctions (GVA), one needs to make sure that the same subset of the search space is explored by each instance of the optimization algorithms. Remember that in GVA one computes the Clarke taxes using the value of the best allocation for several different settings (excluding each bidder in the computation of his tax).

Remark 1 Therefore the same shuffling has to be used for all these instances of optimization. Also the same parameter $T$ has to be used for each instance optimization (at each given total weight of the solution).

At each optimization instance that excludes a bidder one only needs the weight of the optimal solution, and only satisfiable ( P ) is computed there for each given total weight.

### 5.2 Secure Stochastic Search with arithmetic circuits

The secure stochastic algorithms based on mix-nets suffer from the fact that the cost of shuffling remains the same as for the non-stochastic complete approaches. This was particularly negative in the case of SMPC-DisCSP1 where the cost of the shuffling is the main cost.

This problem is reduced in algorithms with shuffling based on arithmetic circuits. Namely, with shuffling based on arithmetic circuits one does not need to compute the whole shuffling. With SMPC-DisCSP4, it is possible to only compute the first $T$ elements of the shuffled problem (see Algorithms 9, 10), and 11).

```
function Shuffle(s,k,r,T)
    for }j=1\mathrm{ to }T\mathrm{ do
        Perm(s,j,r[j],j,k,k);
```

Algorithm 9: Shuffling a vector $s$ with $k$ shared secrets, and a random vector $r$ obtained with Algorithm 2

It can be noted that in secure stochastic algorithms based on arithmetic circuits we succeed to reduce the cost of shuffling and unshuffling from $O\left(k^{2}\right)$ to $O(k T)$ multiplications of secrets. With this improvement the complexity of SMPC-DisCSP4ac

```
function Shuffle(s,k,r,T)
    for j=T to 1 do
        Perm(s,j,r[j],j,k,k);
```

Algorithm 10: Un-shuffling a vector $s$ with $k$ shared secrets, when the shuffling was defined by random secret vector $r$.

```
function SMPC-DisCSP4ac(T,(X,D,C))
    for }i=1\mathrm{ to }k\mathrm{ do
        S[i]= \}\mp@subsup{|}{\phi\inC}{}\phi(\mp@subsup{\epsilon}{i}{})
    R=RandomVector(T);
    SHUFFLE(S,k,R,T) //using the mixnet;
    h[1]=1;
    for i=2 to T do
        h[i]=h[i-1]*(1-S[i-1]);
        S[i]=S[i]*h[i];
    /* S[T]=S[T]*cmp(RS(0,q-1),p*q)// fine tuning*/;
    for }i=T+1\mathrm{ to }k\mathrm{ do
        S[i]=0;
    UNSHUFFLE(S,k,R,T);
    set solution S to 0 with probability p; //optional;
    return S// the solution can be extracted from S as in [Sil05b];
```

Algorithm 11: SMPC-DisCSP4ac, solving a CSP $(X, D, C)$ with $k$ alternatives allowed by the public constraints, and exploring $T$ alternatives.
decreases, but remains high since $k$ is large for hard problems (can be exponential in the problem size).

In conclusion the most appropriate algorithm for Stochastic Search is SMPCDisCSP1 which has polynomial space requirements and whose computational (time) complexity can be bounded to low values being linear in $T$ and in the problem size.

SMPC-DisCSP4ac (with arithmetic circuits) has a time complexity significantly smaller than MPC-DisCSP4 $\left(O(k(T+c))\right.$ versus $\left.O\left(k^{2}\right)\right)$. This implies that the size of the problems solvable with SMPC-DisCSP4' is larger than the size solvable with MPC-DisCSP4, which had the best complexity among complete algorithms.

Remark 2 (SMPC-DisCSP1ac) Arithmetic circuit shuffling for SMPC-DisCSP1 works by separately permuting each domain (with a separate random vector for each of them). The improvement that can be brought is to only compute the permuted constraint elements that are part of the first $T$ tuples.

The shuffling for SMPC-DisCSP1 is not expensive. Therefore possible improvements in versions based on arithmetic circuit shuffling are less significant, not changing the time complexity.

## 6 Optimization Problems

It is known that, in general, Constraint Optimization Problems (COP) are NP-hard. Existing arithmetic circuits for secure protocols solving such problems are exponential in the number of variables, $n$. Using variable elimitation techniques [Dec90] COPs can be solved with computation that is exponential only in the induced-width of the Depth First Search tree (DFS) of the constraint graph [PF05b], i.e. smaller than $n$. We show how to construct an arithmetic circuit with this property and solving any COP. For forest constraint graphs, this leads to a linear cost secure solver.

Combinatorial optimization is an important operation in many problems. One important formalism for modeling combinatorial optimization is the constraint optimization problem (COP). A constraint optimization problem $(X, D, C)$ is defined by a set of variables, $X=\left\{x_{1}, \ldots, x_{m}\right\}$, with domains from $D=\left\{D_{1}, \ldots, D_{m}\right\}$, and a set of weighted constraints $C=\left\{\phi_{0}, \ldots, \phi_{m}\right\}$, each such constraint $\phi_{i}$ specifying a distinct cost associated with each assignment of values to a subset $X_{i}$ of $X$.

An assignment is a pair $\left\langle x_{i}, v\right\rangle$ where $v \in D_{i}$. A solution of the COP is a tuple of assignments $\epsilon$ with values for each variable in $X$ such that the sum of the weights associated by the constraints in $C$ to $\epsilon$ is maximized (minimized). Without loss of generality we assume that by optimal solution we understand the solution with maximal weight. If we denote the projection of a tuple $\epsilon$ on a set of variables $X_{i}$ by $\epsilon_{\mid X_{i}}$, then the solution is:

$$
\underset{\epsilon}{\operatorname{argmax}} \sum_{\phi_{i} \in C} \phi\left(\epsilon_{\mid X_{i}}\right)
$$

A distributed COP (DCOP) arises when some constraints are functions of secrets own by some agents from a set $A=\left\{A_{1}, \ldots, A_{n}\right\}$. Without loss of generality we assume that $\phi_{0}$ is the only public constraint and that $X_{m}$, the set of variables in $\phi_{m}$,
contains besides $x_{m}$ only variables $x_{i}$ with $i<m$. Note that such a formulation can be obtained from any DCOP by building a Depth First Search (DFS) tree, introduced later and combining the constraints such that there remains a single constraint per variable (with his ancestors in the tree).

Our work employs the following secure multi-party computation techniques:

- polynomial secret sharing [Sha79]: Each participant $k$ out of $n$ participants receives $\langle s\rangle_{k}^{t}=s+\sum_{i=1}^{t} a_{i} k^{i}$, where $a_{i}$ is a secret random number. The secret can be reconstructed with the collaboration of $t+1$ participants using $s=\sum_{k=1}^{t+1} l_{k, t}$, where $l_{k, t}$ are the corresponding Lagrange coefficients.
- addition of shared secrets [BOGW88]: $\left\langle s_{1}+s_{2}\right\rangle_{k}^{t}=\left\langle s_{1}\right\rangle_{k}^{t}+\left\langle s_{2}\right\rangle_{k}^{t}$
- resharing shared secrets: To reshare a secret $\langle s\rangle^{t}$ with another threshold $t^{\prime}$, each share $\langle s\rangle_{k}^{t}$ is shared with $\left(t^{\prime}+1, n\right)$-polynomial sharing scheme.
- multiplication of shared secrets [BOGW88]: $\left\langle s_{1} * s_{2}\right\rangle_{k}^{2 t}=\left\langle s_{1}\right\rangle_{k}^{t} *\left\langle s_{2}\right\rangle_{k}^{t}$
- arithmetic circuit evaluation with additive secret sharing [Gol04]: Each participant $k, k>1$ out of $n$ participants receives $[s]_{k}=a_{i}$ where $a_{i}$ is a random number. Participant 1 gets $[s]_{1}=s-\sum_{i=2}^{n} a_{i}$. The secret could be reconstructed with $s=\sum_{k=1}^{n}[s]_{k}$. Addition of additively shared secrets is done with $\left[s_{1}+s_{2}\right]_{k}=\left[s_{1}\right]_{k}+\left[s_{2}\right]_{k}$. Multiplication is done using oblivious transfers [Gol04].
- secure test [DFNT05]: $\delta(x)$ returns 1 if $x=0$ and 0 otherwise.
- secure Kronnecker's $\delta$ [Kil05]: $\delta_{K}(x, y)=\delta(x-y)$ returns 1 if $x=y$ and 0 otherwise.
- secure comparison [DFNT05]: $\operatorname{cmp}(x, y)$ returns 0 if $x<y$ and 1 otherwise.
- secure max: $\max (x, y)=c m p(x, y) *(x-y)+y$.


### 6.1 Optimization Background

DCOPs have been addressed with various techniques that differ both in efficiency and in their privacy guarantees. The former techniques seeking the strongest privacy guarantees are based on secure multiparty computation and scan several times the whole search space, i.e. Cartesian product of domains in $D$, once for each possible total weight [SM04]. An optimization protocol specialized on generalized Vickrey auctions and based on dynamic programming is proposed in [YSO4] and is significantly more efficient, but does not randomize the selection of the solution, needed for reaching the highest level of privacy [SR04]. DPOP, a dynamic programming algorithm for solving (D)COPs was proposed in [PF05b] and consists of a Viterbi-like combination of a maximization and decoding [Vit67]. The algorithm in [PF05b] can also be seen as a clever heuristique for variable elimination [Dec90], or as a parallelization of ADOPT [MTSY03], and is based on a different concept of privacy [SF02].


Figure 3: For a COP with primal graph depicted in (a), two possible DFS trees are (b) with induced width with 2 and (c) with induced width 2.

### 6.1.1 Variable Elimination

Variable Elimination is a principled technique for complexity reduction in COPs. It consists of replacing all the constraints (objective functions) linked to a variable chosen for elimination by the projection of their composition on the remaining variables. A heuristique for selecting the variables to be eliminated next is provided by the DFS tree [PF05b].

### 6.1.2 DFS tree

The primal graph of a COP is the graph having the variables as nodes and having an arc for each pair of variables linked by a constraint [Dec03]. A Depth First Search (DFS) tree associated to a COP is a spanning tree generated by the arcs used for visiting once each node during some depth first traversal of its primal graph. DFS trees were first successfuly used for Distributed Constraint problems in [CDK00]. The property exploited there is that separate branches of the DFS-tree are completely independent once the assignments of common ancestors are decided.

Definition 4 (neighbor nodes) The nodes directly connected to a node in a primal graph are said to be its neighbors.

In Figure 3.a, the neighbors of $x_{3}$ are $\left\{x_{1}, x_{5}, x_{4}\right\}$.
Definition 5 (ancestor nodes) The ancestors of a node are the nodes on the path between it and the root of the DFS tree, inclusively.

In Figure 3.b, the ancestors of $x_{2}$ are $\left\{x_{5}, x_{3}\right\}$, while $x_{3}$ has no ancestors.
Definition 6 (descendants nodes) The descendants of a node are its children in the DFS tree, as well as the children of any other of its descendants.

In Figure 3.c, the descendants of $x_{3}$ are $\left\{x_{1}, x_{4}, x_{2}\right\}$, while $x_{2}$ has no descendants.
A neighboring ancestor of a node is any node that is both a neighbor in the primal graph and an ancestor in the used DFS tree. The induced width of a DFS tree is given by the the number of ancestors that are neighboring him or some of its descendants Two examples of DFS trees for a COP primal graph are shown in Figure 3. It can be noted that trees with different induced widths can be obtained. Several heuristiques had been used in the past for building DFS trees with reduced width, a prefered optimal technique consisting of a branch \& bound procedure.

The ancestors of $x_{i}$ are all the nodes of the path between root of the tree and $x_{i}$, inclusively the root. The descendants of $x_{i}$ are all the nodes for which $x_{i}$ is an ancestor. We use the following notation. Let:

- $F_{x}$ be the parent of $x$ (in Figure 3.b $F_{x_{5}}=x_{3}$ )
- $S_{x}$ be the children of $x$ (in Figure 3.c $S_{x_{3}}=\left\{x_{1}, x_{4}\right\}$ )
- $P_{x}$ be the neighbor ancestors of $x$ (in Figure 3.c $P_{x_{2}}=\left\{x_{1}, x_{5}\right\}$ )
- $G_{x}$ be the induced parents of $x$, i.e., ancestors that are neighbors for $x$ or for some descendant of $x$ (in Figure 3.c $G_{x_{1}}=\left\{x_{3}, x_{5}\right\}$, since $x_{5}$ is the neighboring ancestor of the descendant $x_{2}$ )

In this work we do not address heuristiques for building DFS trees, but consider that such a tree is provided.

### 6.1.3 DFS-based Variable Elimination

A heuristique for selecting the order to eliminate variables based on exploiting the DFS-tree is proposed in [PF05b]. The idea is that before eliminating a node in the DFS tree one should first eliminate its children. In a centralized approach, such an order could be generated by either a postorder traversal or a reversed level-order traversal. This heuristiq guarantees that the arity of the largest constraint that will be added to the problem (and therefore the complexity of the algorithm) is bounded by the distance between two neighbors in the DFS tree. This is bounded by the depth of the tree and potentially much smaller than $n$. The advantage of this heuristique is that the quality of an elimination order can be easily evaluated.

### 6.1.4 Secure Optimization

A secure optimization algorithm for DCOP is proposed in [SM04]. It chooses randomly one of the values with the optimal value and reveals the total weight of the solution and the corresponding assignments only if desired and only to agreed participants. To ensure random selection of the solution, shuffling of values is done prior to solving. The result of the computation will be unshuffled. In [Sil04] it is shown how to make the selection with a uniform random distribution. However, the complete versions of these techniques are always exponential in the size of the search space (as defined by the public constraints).

Remark 3 The fact that a mix-net leads to a random solution only hold for an agent (or a set of agents) if the 'other' agents shuffle all variables.

If at least one agent does not shuffle a variable, this do not hold for the sub-group that shuffle, and the benefit of the shuffling is completely lost

Theorem 1 All agents must be involved in shuffling each variable's domain of a problem.

Proof. If some agent $A_{i}$ does not shuffle a variable $x_{j}$, then all the other agents can make a coalition, find the order on $x_{j}$ with which the problem was solved, and learn a set of values of $x_{j}$ for which there is no optimal solution. Eliminating $x_{j}$ for those values leads to a projection of constraint on the remaining variables, $X^{\prime}$, from which they can infer bounds on $A_{i}$ 's costs on those remaining variables $X^{\prime}$ (lower bounds at minimization respectively upper bounds at maximization). $\quad$ The techniques in [Sil05c] achieve the required type of shuffling/unshuffling.

In the following we show how to formulate arithmetic circuits for securely computing an optimal solution of a DCOP using DFS-based Variable Elimination. As in the non-secure version, the algorithm has two parts, un (upward) dynamic programming step, and a (downward) decoding step.

### 6.2 Arithmetic Circuits

The data structure we employ as well as their usage is depicted graphically in Figure 4. For each node $x_{i}$ there is a separate set of data structures $W_{F_{i}}^{x_{i}}, W_{x_{i}}^{x_{i}}, W_{x_{i}}$. These data structures are accessed by indexing with partial assignment of variables (and are implementable as multi-dimensional matrices).

- $W_{x_{i}}^{x_{i}}$ holds $\phi_{i}$, namely a local costs associated to each assignment of $x_{i}$ and neighboring ancestors of $x_{i}$.
- $W_{x_{i}}$ holds the cummulated cost associated by $\phi_{i}$ and the projection of the constraints of all its children to each assignment of $x_{i}$ and induced parents of $x_{i}$, $G_{x_{i}} \cup\left\{x_{i}\right\}$.
- $W_{F_{i}}^{x_{i}}$ holds the cummulated cost associated by $\phi_{i}$ and the projection of the constraints of all its children to each assignment of induced parents of $x_{i}, G_{x_{i}}$.

On the upward path in the DFS tree, for each node $x_{i}$ one computes for each assignment $S$ of the induced parents $G_{x}$ :

$$
W_{F_{i}}^{x_{i}}[S]=\max _{v \in D_{i}}\left(W_{x_{i}}^{x_{i}}\left[\left\langle x_{i}, v\right\rangle \cup S_{\mid P_{x_{i}}}\right]+\sum_{y \in S_{x_{i}}}\left(W_{x_{i}}^{y}\left[\left(S \cup\left\{\left\langle x_{i}, v\right\rangle\right\}\right)_{\mid G_{y}}\right]\right)\right.
$$

The value $W_{\emptyset}^{r}$, where $r$ is the index of the root node, is the weight of the optimal solution. The steps of computation are detailed in the Algorithm 12.


Figure 4: Data structures in the Upward computation.

### 6.2.1 Unsecure decoding of solution

If we also want to reaveal the assignment with the optimal value, it can be done with the following arithmetic circuit, for the downward path (e.g., level-order traversal from

```
procedure \(\operatorname{Upward}\left(x_{i}\right)\) do
    foreach \(\left(y \in S_{x_{i}}\right)\) do
        Upward (y);
    foreach tuple \(\epsilon\) for \(G_{x_{i}} \cup\left\{x_{i}\right\}\) do
        \(W_{x_{i}}[\epsilon]=W_{x_{i}}^{x_{i}}\left[\epsilon_{\mid\left\{x_{i}\right\} \cup P_{x_{i}}}\right]+\sum_{y \in S_{x_{i}}}\left(W_{x_{i}}^{y}\left[\epsilon_{\mid G_{y}}\right]\right.\);
    foreach tuple \(\epsilon\) for \(G_{x_{i}}\) do
        \(W_{F_{i}}^{x_{i}}[\epsilon]=\max _{v \in D_{i}}\left(W_{x_{i}}\left[\epsilon \cup\left\langle x_{i}, v\right\rangle\right]\right) ;\)
```

Algorithm 12: Arithmetic circuit for the upward (dynamic programming) step. At the first call, the parameter $x_{i}$ is the root, $r$, of the DFS tree.
root).
The value of the whole tree, $V_{r}$, is the shared secret $W_{\emptyset}^{r}$ computed at the upward step. At variable $x_{i}$ with subtree value $V_{x_{i}}$, for each tuple $\epsilon$ with value $W$ ' in the structure $W_{x_{i}}$ at assignments equal to the ones selected at previous levels, compute and reveal $\delta_{K}\left(W, W^{\prime}\right)$. If the result is 1 , the corresponding assignment of $x_{i}$ is selected, and the corresponding values $V_{y}$ in $W_{x_{i}}^{y}$ for each child variable $y \in S_{x_{i}}$ is selected as value of the subtree with root $y$.

The problems with this approach, of revealing the solution, is that the algorithm cannot be used in cases where the solution of the COP is only an intermediary computation, e.g. [Sil05a]. We fix this in the following algorithm.

### 6.2.2 Secure decoding of solution

NOTE that a more complete and accurate version of this subsection appears in [SFP06]! In particular, the version in this subsection does not treat the case where there are several optimal solutions, which is addressed in [SFP06].

The data structures proposed for the secure downward computation (decoding) are again similar for each node. At node $x_{i}$ we store:

- $d_{x_{i}}$ is a shared secret vector of $\left|D_{i}\right|$ boolean values, each indicating the selection of the corresponding value of $D_{i}$. Only one value can be set to true.
- $V_{x_{i}}$ is the shared secret weight associated to the selected optimal solution by the subtree with root $x_{i}$.

After performing the upward computation of dynamic programming, we can decode the solution securely during a preorder or level-order traversal of the tree. On visiting each node $x_{i}$, a procedure is run to compute securely the shared secret assignment of $x_{i}$ using the shared secret assignments of the induced parents, $d_{p}$, for $p \in G_{x_{i}}$, and the shared secret selected weight of this variable, $V_{x_{i}}$. The procedure also computes the inputs for the next recursive procedure calls, at the descendents of $x_{i}$, namely the shared secret selected weight $V_{y}$ of each child $y \in S_{x_{i}}$. The following sums are over all tuples $\epsilon$ of assignments for $G_{x_{i}} \cup\left\{x_{i}\right\}$.

$$
\begin{gathered}
V_{y}=\sum_{\epsilon}\left(\prod_{p \in G_{x_{i}}} d_{p}\left[\epsilon_{\mid p}\right]\right) \delta_{K}\left(V_{x_{i}}, W_{x_{i}}[\epsilon]\right) W_{x_{i}}^{y}\left[\epsilon_{\mid G_{y}}\right] \\
d_{x_{i}}[v]=\sum_{\epsilon_{\mid\left\{x_{i}\right\}}=v}\left(\prod_{p \in G_{x_{i}}} d_{p}\left[\epsilon_{\mid p}\right]\right) \delta_{K}\left(V_{x_{i}}, W_{x_{i}}[\epsilon]\right)
\end{gathered}
$$

```
procedure Downward(CSP) do
    while }\mp@subsup{x}{i}{}\leftarrow\mathrm{ get_InOrder_Next(DFS(CSP)) do
            foreach v\in Di
            d}\mp@subsup{d}{\mp@subsup{x}{i}{}}{}[v]=\mp@subsup{\sum}{\mp@subsup{\epsilon}{{{\mp@subsup{x}{i}{}}}{}=v}{}(\mp@subsup{\prod}{p\in\mp@subsup{G}{\mp@subsup{x}{i}{}}{}}{}\mp@subsup{d}{p}{}[\mp@subsup{\epsilon}{|p}{}])\mp@subsup{\delta}{K}{}(\mp@subsup{V}{\mp@subsup{x}{i}{}}{},\mp@subsup{W}{\mp@subsup{x}{i}{}}{}[\epsilon])
        foreach y}\in\mp@subsup{S}{\mp@subsup{x}{i}{}}{}\mathrm{ do
            Vy= \sum 
```

Algorithm 13: Arithmetic circuit for the downward step.
The computation steps required by the downward phase are detailed in Algorithm 13.

At the end of this computation, the vectors $d_{x_{i}}$ hold a shared unary constraint allowing a single value for $x_{i}$, namely the one in the optimal solution. These unary constraints can then be unshuffled [Sil05c].

### 6.3 Complexity Analysis

The Upward step is called once for each variable $x_{i}$ and the number of operations for each variable is linear in the number of elements of $W_{x_{i}}$, i.e., exponential in $\left|G_{x_{i}}\right|+1$. The total cost for the upward step is $O\left(n d^{g+1}\right)$, where $d$ is the maximum size of a domain of a variable, and $g$ is the maximum value for $\left|G_{x_{i}}\right|$, i.e. the induced width of the used DFS tree.

In the Downward step there exists a while loop for each variable $x_{i}$ and each such cycle has two summations for each element in $W_{x_{i}}$, each term having $\left|G_{x_{i}}\right|+1$ multiplications. The total cost for the downward step is $O\left(n g d^{g+1}\right)$.

Therefore, the total complexity of the secure version is $O\left(n g d^{g+1}\right)$. If the downward version with immediate revelation of assignments in solutions is used, then the complexity is only $O\left(n d^{g+1}\right)$. If shuffling of constraints and unshuffling of solution vectors $d_{x}$ are used to randomize the selection of the solution, then the cost of the shuffling is also added [Sil05c].

### 6.4 Extensions and Applications

A version of Secure Stochastic Optimization can be obtained by trimming the intermediary data structures at each node on the Upward phase. Something similar was proposed for non-cryptographic techniques in [PF05a].

Remark 4 Note that sorting the set of weigths before trimming (in a beam-search like approach) is possible, but only for some applications. E.g., in auctions one has to ensure that the same tuples survive the prunning at each different optimization subproblem appearing in the Clarke tax computations [Sil05a].

An immediate application for secure DCOPs is in performing the intermediary optimizations steps for Clarke tax in generalized Vickrey auctions, and related auction clearance mechanisms. If a secure stochastic version is used, then one has to use the same prunning (with the same surviving tuples) for each different optimization task on a given problem, as explained in [Sil05a].

## 7 Conclusions

In this work we have proposed a new family of secure solvers for distributed Constraint Satisfaction Problems (disCSPs). While most existing techniques were complete and inapplicable to large instances, the new techniques can be used to address large problems.

We have proposed stochastic versions for each of the complete secure multi-party algorithms MPC-DisCSP1 and MPC-DisCSP4, based on shuffling with mixnets or with arithmetic circuit. MPC-DisCSP1 is remarkable for its polynomial space requirements while MPC-DisCSP4 for its low time complexity and for the uniform distribution in selecting solutions.

The new versions only explore a subset of the search space of the problem, subset whose size is specified as a parameter. We have thus analyzed in detail three newly obtained versions: SMPC-DisCSP1, SMPC-DisCSP4, and SMPC-DisCSP4ac.

As its complete counterpart, SMPC-DisCSP1 requires only polynomial space. Unexpectedly, the versions obtained from MPC-DisCSP4 are much less appropriate for addressing large problems, but maintain the desirable property of selecting solutions with a uniform distribution. Among SMPC-DisCSP4 and SMPC-DisCSP4ac, the latter (based on arithmetic circuits) presents the largest speed-up in comparison to its complete version. The algorithm of choice for tackling large problems are therefore the ones based on MPC-DisCSP1 (SMPC-DisCSP1 and SMPC-DisCSP1ac), and their time complexity is linear in the problem size and in a parameter deciding the size of the explored search space.

## References

[BOGW88] M. Ben-Or, S. Goldwasser, and A. Widgerson. Completeness theorems for non-cryptographic fault-tolerant distributed computating. In STOC, pages 1-10, 1988.
[CDK00] Z. Collin, R. Dechter, and S. Katz. Self-stabilizing distributed constraint satisfaction. Chicago Journal of Theoretical Computer Science, 2000.
[Dec90] R. Dechter. Enhancement schemes for constraint processing: Backjumping, learning, and cutset decomposition. AI’90, 1990.
[Dec03] Rina Dechter. Constraint Programming. Morgan Kaufman, 2003.
[DFNT05] I. Damgård, M. Fitzi, J. B. Nielsen, and T. Toft. How to split a shared number into bits in constant round and unconditionally secure. Cryptology ePrint Archive, Report 2005/140, 2005. http://eprint.iacr. org.
[FMW01] E.C. Freuder, M. Minca, and R.J. Wallace. Privacy/efficiency tradeoffs in distributed meeting scheduling by constraint-based agents. In Proc. IJCAI $D C R$, pages 63-72, 2001.
[Gol04] O. Goldreich. Foundations of Cryptography, volume 2. Cambridge, 2004.
$\left[\mathrm{HCN}^{+} 01\right]$ T. Herlea, J. Claessens, G. Neven, F. Piessens, B. Preneel, and B. Decker. On securely scheduling a meeting. In Proc. of IFIP SEC, pages 183-198, 2001.
[Kil88] J. Kilian. Founding cryptography on oblivious transfer. In Proc. of ACM Symposium on Theory of Computing, pages 20-31, 1988.
[Kil05] E. Kiltz. Unconditionally secure constant round multi-party computation for equality, comparison, bits and exponentiation. Cryptology ePrint Archive, Report 2005/066, 2005. http://eprint.iacr.org.
[MTSY03] P.J. Modi, M. Tambe, W.-M. Shen, and M. Yokoo. An asynchronous complete method for distributed constraint optimization. In AAMAS, Melbourne, 2003.
[NZ05] K. Nissim and R. Zivan. Secure discsp protocols - from centralized towards distributed solutions. In DCR05 Workshop, 2005.
[PF05a] Adrian Petcu and Boi Faltings. Approximations in distributed optimization. In Principles and Practice of Constraint Programming CP 2005, 2005.
[PF05b] Adrian Petcu and Boi Faltings. A scalable method for multiagent constraint optimization. In IJCAI, 2005.
[SF02] M.-C. Silaghi and B. Faltings. A comparison of DisCSP algorithms with respect to privacy. In AAMAS-DCR, 2002.
[SFP06] M.-C. Silaghi, B. Faltings, and A. Petcu. Secure combinatorial optimization using dfs-based variable elimination. In Symposium on AI and Maths, January 2006.
[Sha79] A. Shamir. How to share a secret. Comm. of the ACM, 22:612-613, 1979.
[Sil02] M.-C. Silaghi. Asynchronously Solving Distributed Problems with Privacy Requirements. PhD Thesis 2601, (EPFL), June 27, 2002. http://www.cs.fit.edu/~msilaghi/teza.
[Sil03] M.-C. Silaghi. Solving a distributed CSP with cryptographic multi-party computations, without revealing constraints and without involving trusted servers. In IJCAI-DCR, 2003.
[Sil04] M.-C. Silaghi. Meeting scheduling system guaranteeing n/2-privacy and resistant to statistical analysis (applicable to any DisCSP). In $3 r d$ IC on Web Intelligence, pages 711-715, 2004.
[Sil05a] M. Silaghi. Using secure discsp solvers for generalized vickrey auctions, complete and stochastic secure techniques. In IJCAIO5 DCR Workshop, 2005.
[Sil05b] M.-C. Silaghi. Hiding absence of solution for a discsp. In FLAIRS'05, 2005.
[Sil05c] M.-C. Silaghi. Zero-knowledge proofs for mix-nets of secret shares and a version of ElGamal with modular homomorphism. Cryptology ePrint Archive, Report 2005/079, 2005. http://eprint.iacr.org.
[SM04] M.-C. Silaghi and D. Mitra. Distributed constraint satisfaction and optimization with privacy enforcement. In 3rd IC on Intelligent Agent Technology, pages 531-535, 2004.
[SPF05] M.C. Silaghi, A. Petcu, and B. Faltings. Secure combinatorial optimization using dfs-based variable eliminations. Technical Report CS-2005-15, FIT, 2005.
[SR04] M.-C. Silaghi and V. Rajeshirke. The effect of policies for selecting the solution of a DisCSP on privacy loss. In AAMAS, pages 1396-1397, 2004.
[Vit67] A.J. Viterbi. Error bounds for convolutional codes and an asymtotically opti mum decoding algorithm. IEEE Trans. on Information Theory, 13(2):260-267, 1967.
[Yao82] A. Yao. Protocols for secure computations. In FOCS, pages 160-164, 1982.
[YS04] M. Yokoo and K. Suzuki. Generalized Vickrey Auctions without ThirdParty Servers. In FC04, 2004.
[YSH02a] M. Yokoo, K. Suzuki, and K. Hirayama. Secure distributed constraint satisfaction: Reaching agreement without revealing private information. In Proc. of the AAMAS-02 DCR Workshop, Bologna, July 2002.
[YSH02b] M. Yokoo, K. Suzuki, and K. Hirayama. Secure distributed constraint satisfaction: Reaching agreement without revealing private information. In $C P, 2002$.


[^0]:    *The last section, on Optimization Problems, was added during a September 2005 update, and contains material from [SPF05]. A subsequent more complete description of that section appears in [SFP06]. The section on the new privacy concept from [Sil05a], is further extended on Dec 31, 2006.

[^1]:    ${ }^{1}$ With DisCSPs[Sil04].

