Extending linearly independent vectors to a basis of the lattice

Consider a lattice $L \subseteq \mathbb{R}^m$ with a basis $B = \{u_1, \ldots, u_n\}$.

1. Let $v = \sum_{i=1}^{n} z_i u_i$ be a vector of L. Show that there exists a basis of L containing v if and only if $\mathrm{GCD}(z_1, \ldots, z_n) = 1$.

2. Suggest an algorithm which takes a basis B of $L \subseteq \mathbb{Z}^m$ and $v \in L$ on the input and decides whether L has a basis containing v. If such a basis exists the algorithm also outputs one such basis.

3. (elective) Suggest an algorithm which takes a basis B of $L \subseteq \mathbb{Z}^m$ and a linearly independent set of vectors $X \subseteq L$ on the input and decides whether L has a basis containing X. If it is the case, the algorithm also outputs such a basis.

Experiments with attacks against the Merkle-Hellman scheme

Recall the basic description of the Merkle-Hellman cryptosystem introduced in 1978:

Let $s_1 < s_2 < \cdots < s_\ell$ be a superincreasing sequence of positive integers, that is, $s_i > \sum_{j=1}^{i-1} s_j$ for any $2 \leq i \leq \ell$. Let N be an integer $N > \sum_{i=1}^{\ell} s_i$ and $c, d \in \mathbb{N}$ such that $cd \equiv 1 \pmod{N}$. For each $1 \leq i \leq \ell$ set $w_i := s_i c \bmod N$.

In the scheme, public key is the sequence $w_1, w_2, \ldots, w_\ell$ and the private key consists of N and d (it may also include $s_1, \ldots, s_\ell$).

Assume someone (typically called Bob) wants to transmit a message which is a sequence of ℓ bits, say $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_\ell \in \{0, 1\}$. Bob computes an integer $C = \sum_{i=1}^{\ell} \varepsilon_i w_i$, an encryption of his message.

In order to decrypt the message from C the following steps are done:

1. Compute $S = Cd \bmod N$
2. For each $1 \leq i \leq \ell$ compute $s_i = w_i d \bmod N$ (this step can be skipped if the sequence $s_1, \ldots, s_\ell$ is a part of the private key).
3. Find $\varepsilon'_1, \ldots, \varepsilon'_\ell \in \{0, 1\}$ such that $S = \sum_{i=1}^{\ell} \varepsilon'_i s_i$ (this can be done easily since $s_1, \ldots, s_\ell$ is a superincreasing sequence).
4. return $\varepsilon'_1, \ldots, \varepsilon'_\ell$

Recall that the correctness of the decryption is based on the following congruence

$$S \equiv Cd \equiv \sum_{i=1}^{\ell} \varepsilon_i (w_i d) \equiv \sum_{i=1}^{\ell} \varepsilon_i s_i (\text{mod } N)$$

Since S and $\sum_{i=1}^{\ell} \varepsilon_i s_i$ are in $\{0, \ldots, N-1\}$, $S = \sum_{i=1}^{\ell} \varepsilon_i s_i$ and $\varepsilon_i = \varepsilon'_i$ for every $1 \leq i \leq \ell$.

Remark: The security of the scheme is based on the hardness of the subset sum problem for the public key, i.e., it should be hard to compute $\varepsilon_1, \ldots, \varepsilon_\ell$ just from $w_1, \ldots, w_\ell$ and C. However, various attacks against this scheme appeared. We consider those introduced in the lecture.

When making the experiments you can use implementation of LLL, nearest plane algorithm and an SVP solver from available libraries, but be aware that these implementations can be different from the algorithms introduced in the lecture.

1. Write an algorithm which constructs public key in the described scheme. Choices $s_1, \ldots, s_\ell, N, c$ should be sufficiently randomized but on the other hand the length of these parameters should be reasonable. Also c should be sufficiently large to hide the original sequence $s_1, \ldots, s_\ell$.

2. Implement the following attack: Given $w_1, \ldots, w_\ell$ the public key and a ciphertext C, that is, there are $\varepsilon_1, \ldots, \varepsilon_\ell \in \{0, 1\}$ such that $C = \sum_{i=1}^{\ell} \varepsilon_i w_i$. In $\mathbb{Z}^{\ell+1}$ consider the lattice with basis $(1, 0 \ldots, 0, w_1)^T$, $(0, 1, 0, \ldots, 0, w_2)^T$, $\ldots$,$(0, \ldots, 0, 1, w_\ell)^T$,$(0, \ldots, 0, C)^T$. Compute an LLL-reduced basis of L and

consider the attack as successful if the basis contains a vector proportional to $(\varepsilon_1, \ldots, \varepsilon_\ell, 0)^T$. Make an experiment giving an idea how is the probability of success of this attack related to ℓ. Regarding the range of ℓ, start for example with $\ell = 5$ and gradually increase its value until the probability of success becomes too small (for example less than 0.01) or the computation of the LLL reduction starts to become too slow.

3. Consider the following modification of the attack from part 2: Instead of LLL use some algorithm which finds a shortest nonzero vector of L. The attack will be considered successful if $\lambda_1(L) = ||(\varepsilon_1, \ldots, \varepsilon_\ell)^T||$. Make an experiment giving an idea how is the probability of success of this attack related to ℓ.

4. Implement the following attack: Given $w_1, \ldots, w_\ell$ the public key and a ciphertext C, that is, there are $\varepsilon_1, \ldots, \varepsilon_\ell \in \{0, 1\}$ such that $C = \sum_{i=1}^{\ell} \varepsilon_i w_i$. In $\mathbb{Z}^{\ell+1}$ consider the lattice with basis $(2, 0 \ldots, 0, w_1)^T$, $(0, 2, 0, \ldots, 0, w_2)^T$, $\ldots$,$(0, \ldots, 0, 2, w_\ell)^T$ and a vector $t = (1, \ldots, 1, C)^T$. Use the nearest plane algorithm to find a vector $y \in L$ with the property

$$||y - t|| \leq 2^{\frac{\ell}{2}} \varrho(t, L)\,.$$

We consider this attack successful if $\varepsilon_1, \ldots, \varepsilon_\ell$ can be found looking at coordinates of $y - t$ (see the argument in the proof of NP-completeness of CVP). Once again make an experiment giving an idea how is the probability of success of this attack related to ℓ.

5. Now forget about Merkle-Hellman scheme and consider a related problem: Assume that $w_1, \ldots, w_\ell$ are randomly chosen integers from interval $\{2^{\ell^2}, \ldots, 2^{\ell^2+1}\}$ and S is an integer of the form $S = \sum_{i=1}^{\ell} \varepsilon_i w_i$, $\varepsilon_i \in \{0, 1\}$. For these $w_1, \ldots, w_\ell, S$ we want to find $\eta_1, \ldots, \eta_\ell \in \{0, 1\}$ such that $\sum_{i=1}^{\ell} \eta_i w_i = S$. Adapt the method from 2. and write an algorithm which could give an answer to this problem using LLL reduction. Make an experiment giving an idea how is the probability of success of this algorithm related to ℓ. Is it possible to use observation from this experiment to increase the performance of the attack from part 2?

6. (elective) One can increase security of the scheme adding more rounds of hiding initial sequence $s_1, \ldots, s_\ell$. In the scenario described above when having $w_1, \ldots, w_\ell$ we choose N', c', d' such that $N' > \sum_{i=1}^{\ell} w_i$, $c'd' \equiv 1 \pmod{N'}$. For $1 \leq i \leq \ell$ let $w'_i = w_i c' \bmod N'$. The public key is $w'_1, \ldots, w'_\ell$. Implement the attack from 2. against this modified scheme and decide whether this modification increases resistance of the scheme against this attack. You can also increase the number of rounds.

Remark: You can find information about a theoretical background behind some of these experiments in paper *The Rise and Fall of Knapsack Cryptosystems* by Andrew Odlyzko.