# A Lightweight Lattice-based Homomorphic Privacy-Preserving Data Aggregation Scheme for Smart Grid

报告人：徐梓桑

# 本文摘要

- 本文主要提出了一种轻量级的基于网格（Lattice-based）的同态密码系统，其用于保护用户的隐私以及信息融合。应用背景是智能电网。
- 该文提出的方案可以再不涉及智能电表的情况下对其读数进行拟合。智能电表或者其中间基站不能对这些拟合数据进行解密，但可以验证其真实性。
- 本文采用的加密方案只基于简单的算术运算，能减少智能应用的计算负担。

# 智能电器分类

- 根据文献[3]，按照智能家用电器的通信需求，将其分为四类：
  - 第一类为小负载，例如灯泡，充电器等。只有当其连接时才需要通知控制中心。
  - 第二类为不可控制的负载。例如，炉灶。其根据用户的需求而工作，并且不能容忍过长的延时。这类电器需要最少的通信基础设施，只发送其功率以及与其使用时间。
  - 第三类为可控的大负载，例如，空调和洗衣机。这类电器通过智能电表向控制中心发送例如预期负载，预期使用时间等。这些电器只有在收到验证正确的信息后才能运行。控制中心可以根据用户的设定，动态定价等来接受或拒绝请求。
  - 第四类是通过通信来调度其充电过程的电动汽车。由于其对通信需求量很大，因此将其单独分为一组。

# 目前研究方向

- 目前对智能电网安全性的研究可分为三个部分：
- 第一，将智能电表连接到硬件设备，例如：temper-resistance devices或者电池，来掩盖该区域的真正用电消耗。虽然这种方法减轻了计算和通信的负担，但该做法十分昂贵并且需要经常进行维护。
- 第二，通过在家庭局域网中人为添加噪声，扰乱电量消耗的读数，再在控制中心滤除噪声。该方法能节省智能电器的计算能力，但却无法精确的重构出其发送的原始信息。
- 第三，采取加密方案。例如ABE加密等。其又可细分为：1，使用认证方案，2，使用匿名方案，隐藏智能电表与其消费信息的联系。两种方案都可保证信息的完整性和机密性，但都会消耗计算和通信资源。

# 本文系统模型

- 1. 网络模型
- 每个居民区由一个主要控制中心提供服务，而一个控制中心连接多个基站（BS），则BSs={BS1，BS2，……，BSh}。
- 每一个基站负载该区域的一组HAN，则HANs={HAN1，HAN2，……，HANm}。HAN可为一栋房屋或一个单元。
- 每个HAN都有一个智能电表SM，其连接着所有家用智能电器APs={AP1，AP2，……，APn}，每个AP都可不通过SM直接与其他AP进行通信。
- HAN中的通信方式采用短距离通信方式，例如，蓝牙和ZigBee。HAN和BS之间通过WIFI通信。控制中心CC，BS，和SM都有一个由独立的可信机构（TA）提供的公共密匙。每个AP都有一个由TA提供的唯一ID号，并存储在安全的存储器中。模型如图1所示

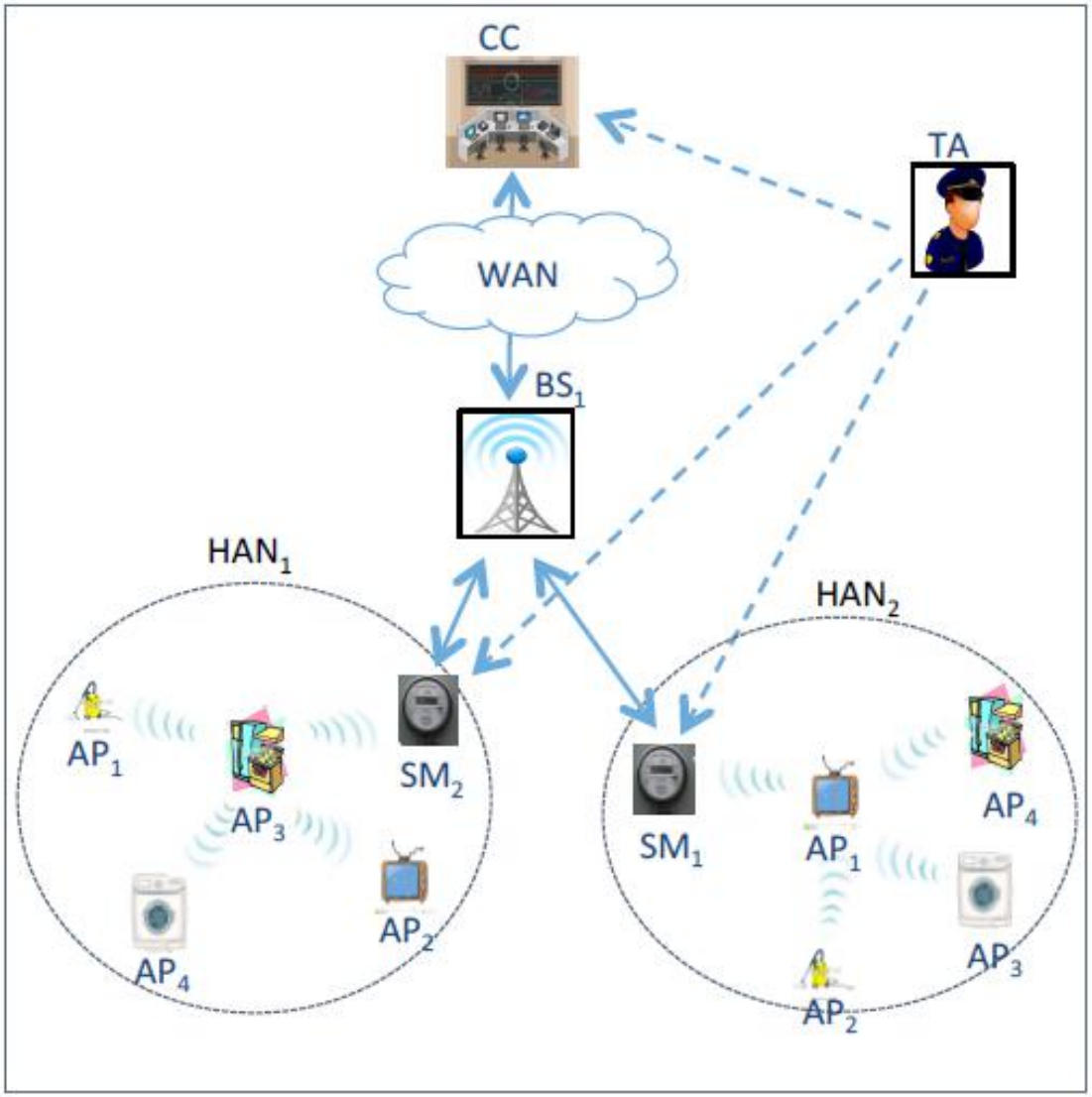



Fig. 1: System Model.

2. 攻击者模式和安全需求

- 本文认为攻击者A可以窃取到AP与SM之间以及SM到BS之间，BS到CC间的转发消息。A可以进行一些主动攻击，例如：伪造捕获的数据或进行重放攻击。因此，我们需要保证以下两点来阻止A的恶意攻击：
- 用户的隐私：确保任何攻击者都无法获得有关于HAN的消费的信息。此外，CC不应该获得该地区每个用户的消费模式。
- 真实性和数据完整性：保证真实性和保密性，使得A即使拦截到一个消息也无法解析出信息。保证数据完整性，使得假设A重发或者修改了某个消息，也能被检测到。

3. 设计目标：

- 满足网络安全需求；保证消费者隐私；防止对设备或消息的任何非法访问或修改，同时还在通信和计算开销方面做到轻量化。

# Preliminaries

- 基于晶格的同态加密方案
- 本文利用向量空间结构将消息加密为噪声晶格。因此，其能以低计算复杂性保证消息的安全性，因为其主要是在向量空间中执行简单的加法和乘法运算。

- 密钥生成：
- 该方案定义了五个全局整形参数：
- N为明文向量坐标数目。r是构成它们的环的特性。l是可以完成的同态运算的最大数量，n是公钥中的软扰动矩阵的数量，以及εmax是用于插入噪声的随机向量的坐标的上限。

- 密钥生成加密解密见论文

# 本文提出的方案

- 本文提出的方案有两个阶段，第一为初始化阶段，通过SM和BS建立AP和CC之间的安全连接，第二为，读数聚合阶段，其组织耗电读数的汇总操作。

1.初始化阶段

- 首先由可信机构TA指派一些公共密钥给CC,BS,SM。

- 对于CC来说，其公共密钥为{Mcc0，Mcc1，……Mccn}，其中Mcc0为硬噪声矩阵，其余的都为软噪声矩阵。AP用这些密钥来加密其读数。CC的私有密钥为 $\mathscr{P}_{cc}(\cdot), M_{cc}, \triangle_{cc}$.
- 对于每个BS，公共密钥为{Mbs0，Mbs1，……Mbsa}，其中Mbs0为硬噪声矩阵，其余的都为软噪声矩阵。其私有密钥为 $\mathscr{P}_{bs}(\cdot), M_{bs}, \triangle_{bs}$.

- 对于每个SM，公共密钥为{Msm0，Msm1，……Msmi}，其中Mbs0为硬噪声矩阵，其余的都为软噪声矩阵。其私有密钥为 $\mathscr{P}_{sm}(.), M_{sm}, \triangle_{sm}$.

- 每个AP都有一个TA指定的唯一ID，{AP1，……APm}，m为HAN中的AP总数。根据ID排列一个固定的顺序，使得每个AP轮流成为每轮的拟合器。
- 每个AP在安全区域存储一个加密的ID，加密方式为

$$ID_{j-enc} = ID_j * M_{sm0} + \sum r_i * M_{smi} \qquad (8)$$

该ID由SM的公共密钥加密，因为AP在作为聚合器聚合信息时，需要提供其身份信息给SM。

2.读取聚合数据阶段

- 见论文

3.控制信息

- 如果来自组3或组4的任何AP需要向CC发送请求，例如改变其负载或使用持续时间，则可以通过SM和BS直接向CC发送请求，并且不等待新的拟合。这些消息称为控制消息。
- APj发送控制信息zj给SM，SM对其签名并通过BS发送给CC。

# 安全分析

隐私：

- 所有的数据都是被加密的，且每轮的拟合器都不同。
- 本文认为A是不能获得AP的ID，若要获得这些ID，就需要物理访问AP，即使获得了，也只能得到这一个AP的数据。这对于SM也是一样的。
- 而BS收到的是整个小区的电力消耗，且整个消息是加密的，所以BS也不能提取出关于家庭用电的消耗情况。CC是整个城区的总体消费总量，因此不能提取出特定对象的实时消费模式的数据。

真实性和消息的保密性和完整性：

- 只有授权方才能访问消息的内容。
- AP只接收被加密的读数，不发送。
- HAN拟合后的数据只有CC能解密，AP和SM都不行。
- BS中也不能从整个区域的总消耗中提取出任何数据。
- 即使SM或BS被妥协了，攻击方也无法解析消息，因为SM和BS没有解密密钥。因此，窃听攻击都是无效的。
- A不能伪造HAN到BS的传输消息，因为A没有SM和BS的私有密钥。而由于附加的时间戳和随机数值，重放攻击也无法成功。

- 而本加密方法的安全性与隐藏晶格的难度问题息息相关（hardness of hidden lattice problem，HLP），其是基于以某种方式打乱原晶格的次序，使得第三方无法从打乱的晶格恢复为原始晶格。

# 性能评估

通信开销

- 每轮数据，每个AP都只发送一次，SM,BS,也只发一次，所以很小。
- 对于控制信号来说，只用组3和组4的才需要发送其请求给CC，这些信息直接发送给SM，再通过BS给CC。
- 本文假设每个HAN最大有三个这样的AP，即三个控制消息。这些消息偶尔发送。假设每个AP需要每天发送1-2个控制消息，则一天最多6个控制消息。开销很小。

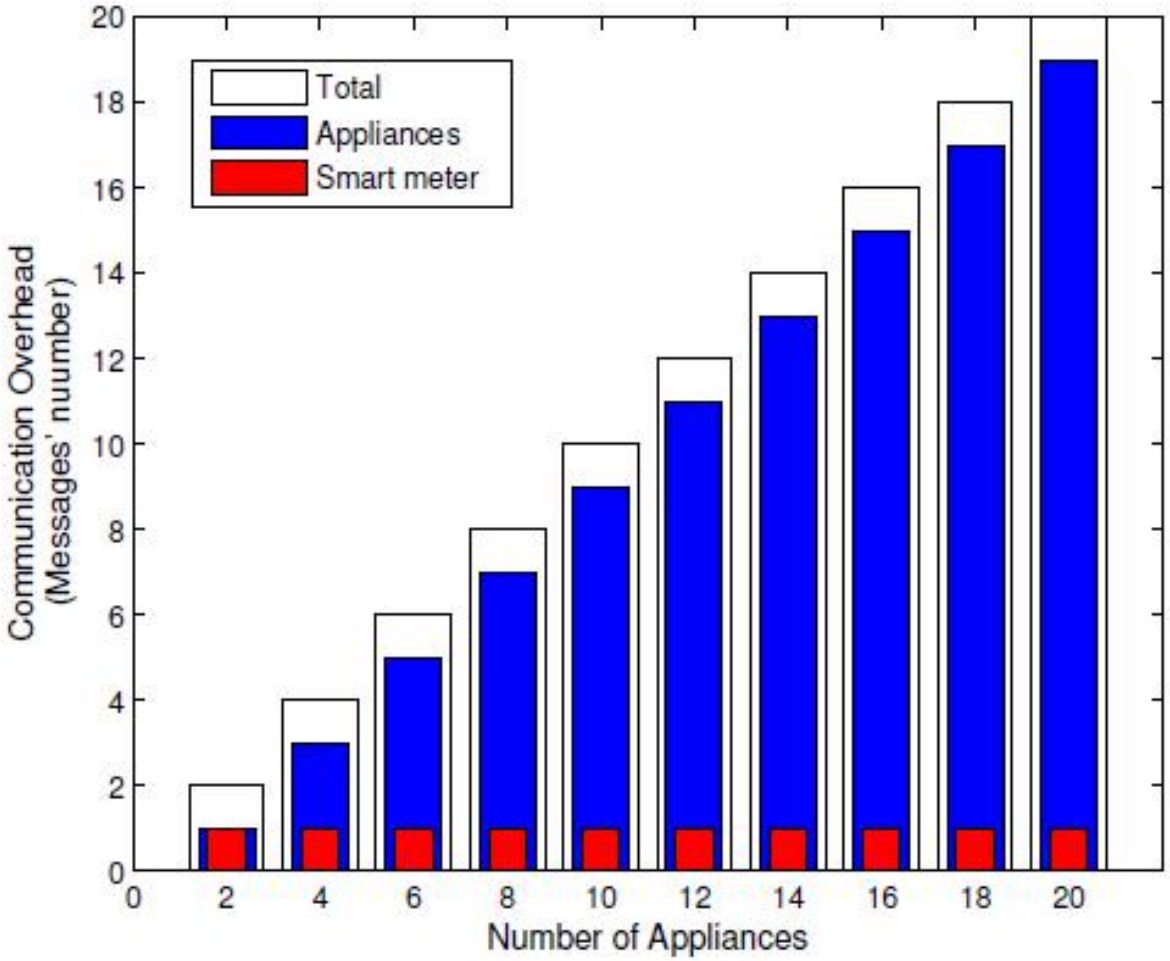



Fig. 2: Communication Overhead per Reading Round.

计算开销

- 除了拟合器APs，HAN中的AP加密操作的总次数为n-1，n为HAN中AP的数量，APs做求和运算。
- 每轮下来，SM都对收到的拟合信息签名一次。
- BS拟合接收到的信息，对其结果签名，然后发送到CC。

- 假设在BS覆盖区域内的HAN数量为m。则该区域的总计算消耗为 $[m*([(n-1)*T_e]+T_s+T_v)]+T_s+T_v+T_d$ ，Te是一个加密过程的计算时间，Td是一个解密过程的时间，Ts是一个签名过程的时间，Tv是一个验证过程的时间。
- 控制信息被认为每天6个，本文认为每个HAN最多3个这类AP，每个AP每天最多发2两个命令。

TABLE I: The number of operations for smart devices.

| Number of Operations | Per Round | Per Day |
|---|---|---|
| $APs(Group-1\&2)$ | $1*T_e$ | $h*T_e$ |
| $APs(Group-3\&4)$ | $1*T_e$ | $(h+2)*T_e$ |
| $SM$ | $1*T_s$ | $(h+6)*T_s$ |

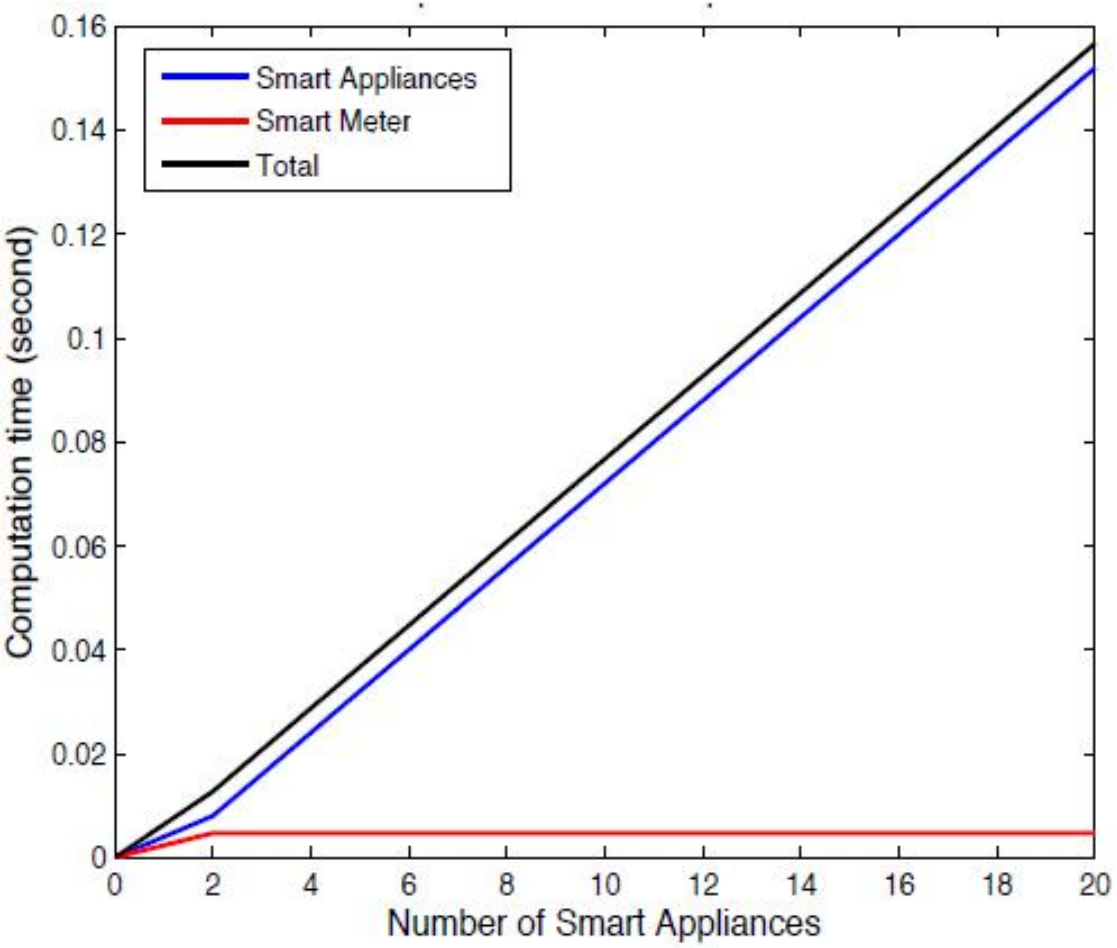
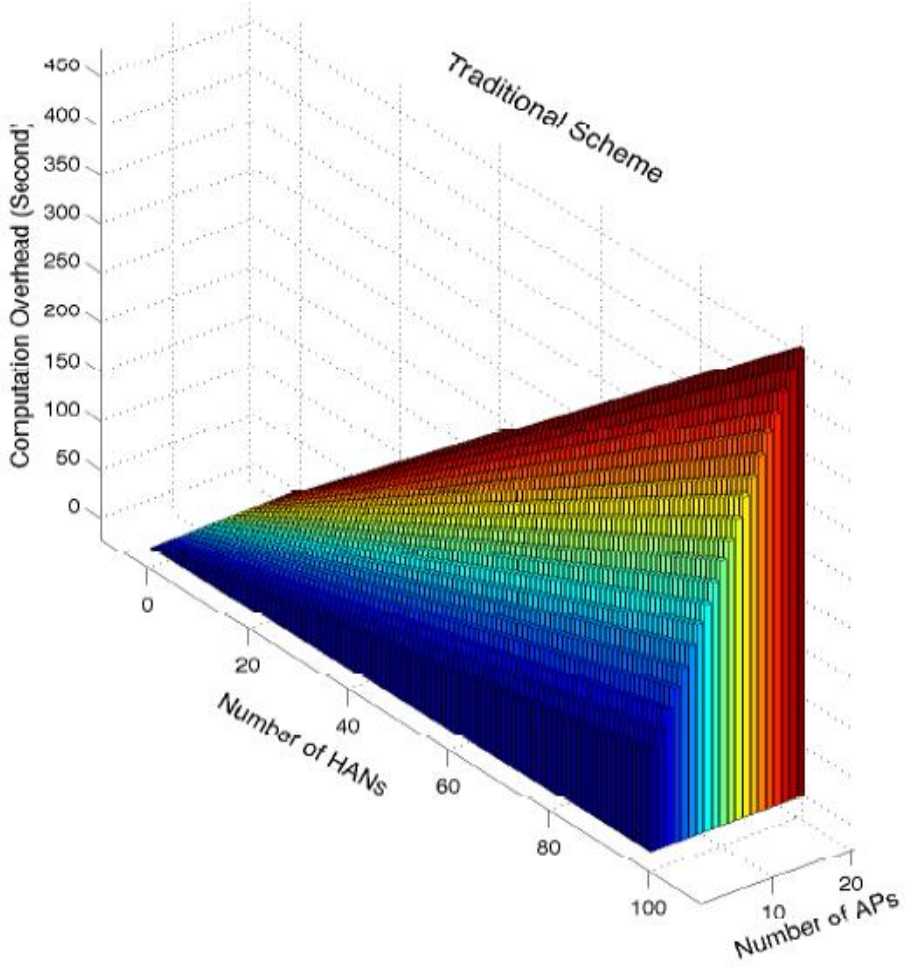



Fig. 4: Computation Delay per Reading Round.

本文最终将本方案与同态 Paillier-based方案进行比较。



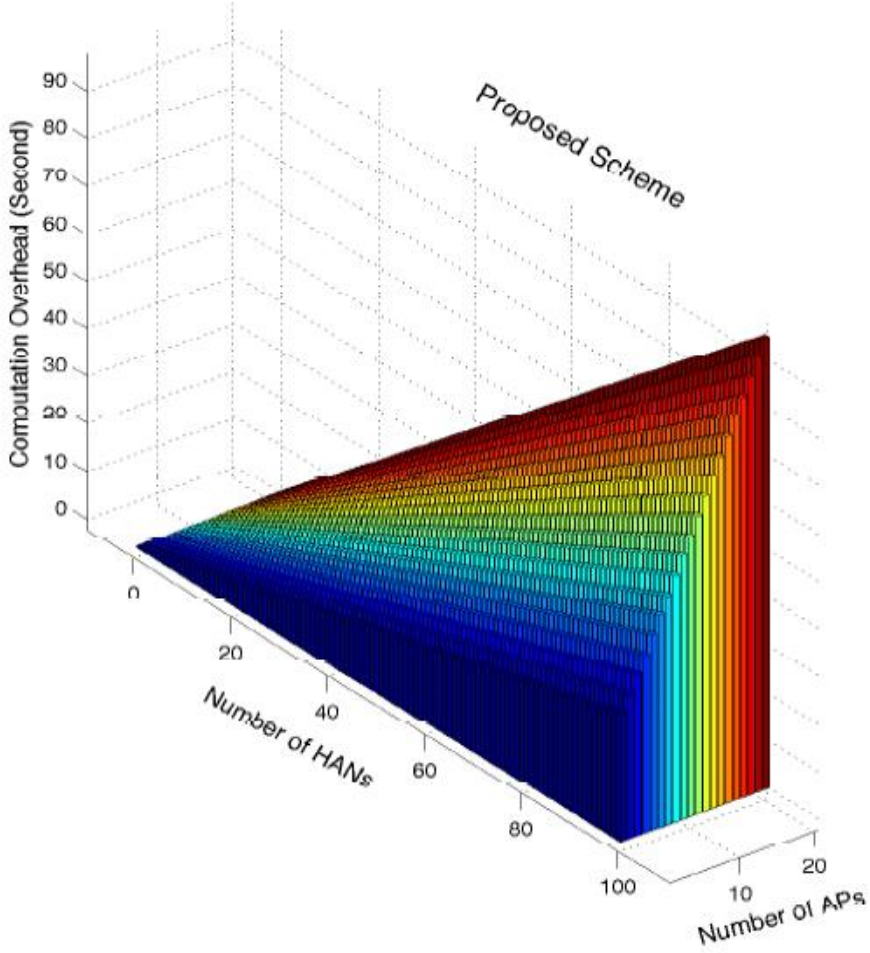




Fig. 7: Computation Delay per Day.