

Data Aggregation of Wireless Sensor Networks Using Artificial Neural Networks *

SUN Lingyi¹, HUANG Xianxiang^{1*}, CAI Wei¹, XIA Meini²

(1. The Second Artillery Engineering Institute, Xi'an 710025, China;
2. China Academy of Space Technology(Xi'an), Xi'an 710100, China)

Abstract: To reduce communication traffic and save energy for wireless sensor networks (WSNs), BPNDAs, a data aggregation algorithm based on back-propagation networks, was proposed, which integrates a three-layer BP neural network with clustering routing protocol. The input layer neuron is located in cluster members, while the hidden layer neuron and the output layer neuron are located in cluster head. Only the processed data represented the features of the raw collected data will be transmitted to the sink, so the efficiency of data gathering is improved and the lifetime of the network is prolonged. Simulation results show that compared with LEACH, the BPNDAs algorithm effectively reduced the data traffic and decreased the energy dissipated of nodes.

Key words: wireless sensor networks; data aggregation; artificial neural networks; cluster

EEACC: 6150P; 6210C

doi: 10.3969/j.issn.1004-1699.2011.01.025

基于神经网络的无线传感器网络数据融合算法 *

孙凌逸^{1*}, 黄先祥^{1*}, 蔡伟¹, 夏梅尼²

(1. 第二炮兵工程学院 202 室, 西安 710025;
2. 中国空间技术研究院西安分院, 西安 710100)

摘要: 为减少无线传感器网络的通信量,降低能量消耗,设计了一种基于神经网络的数据融合算法(BPNDAs),该算法将BP神经网络和传感器网络分簇路由协议有机结合,将每个簇设计成一个神经网络模型,通过神经网络提取原始数据中的少量特征数据,然后将特征数据发送给汇聚节点,从而提高数据收集效率,延长网络生存时间。仿真实验证明,与LEACH算法相比,该算法可有效减少网络通信量,降低节点能耗。

关键词: 无线传感器网络;数据融合;神经网络;分簇

中图分类号: TP393

文献标识码: A

文章编号: 1004-1699(2011)01-0122-06

无线传感器网络(WSNs)是一种新的信息收集(或事件探测)的范式,它依靠众多传感器节点协作地感知、采集和处理网络覆盖区域中的对象信息,并发送给观测者。在该网络中,部署在远程环境中的传感器节点在没有任何网络拓扑先验信息的情况下完成自行配置,其最终目标是监测传感器区域中感兴趣的特定事件。由于传感器节点的检测范围往往重叠,同样的事件通常是由众多的传感器节点报道,这就导致了数据冗余^[1]。

同时,由于无线传感网络部署的环境条件可能会干扰传感器读数,甚至破坏传感器节点,传感器测量结果的准确度可能低于预期值,并且覆盖范围也

可能缩小。因此,在无线传感器网络的应用中必须配置相当数量的冗余节点,用以克服节点失效和量测的不准确性。而传感器节点的能量、存储空间与计算能力有限,冗余数据的传送在一定程度上将消耗过多的能量,缩短整个网络的生存期。

再者,无线传感器网络连接多类传感器,如机械、热、化学、光学以及磁传感器等,用以完成对目标的监测或对周围环境的感知。但单个传感器节点只能完成局部环境的监测或感知,而无线传感器网络关注的则是整个网络感知结果的综合。因此,从应用层面上来讲,数据融合在无线传感器网络中也是必需的。在某种意义上,对多源信息进

行数据融合也将产生一个好于单信息源输出的结果^[2]。

针对上述背景, 本文提出了一种基于神经网络的无线传感器网络数据融合算法 (Back-Propagation Networks Data Aggregation, 简称 BPNDA)。BPNDA 数据融合模型以无线传感器网络中普遍采用的分簇路由协议 LEACH^[3] 为基础, 在簇首节点利用 BP 神经网络对簇成员节点采集的原始数据进行特征提取, 然后将代表原始数据的少量特征值发送给汇聚节点, 以达到减少节点数据通信量、节省能量开销和提高信息收集准确度的目标。

1 相关工作

神经网络是由大量的、简单的处理单元 (称为神经元) 广泛地互相连接而形成的复杂网络系统, 它反映了人脑功能的许多基本特征, 能够模拟人的大脑活动, 具有极强的非线性逼近、分布式存储、大规模并行处理、自训练学习、自组织和容错能力等优点, 特别适合处理需要同时考虑诸多因素和条件的、不精确和模糊的信息处理问题。

无线传感器网络和神经网络有特别相似的地方: 无线传感器网络的节点就好比神经元, 具有感受和处理的功能; 而无线传感器网络节点之间的连接, 则相当于连接神经元的突触, 完成信号的传递。无线传感器网络的数据融合与神经网络具有一个共同的基本特征, 即通过对大量的数据进行一定的运算和处理, 得到能够反映这些数据特征的结论性的结果。因此, 可以利用神经网络的方法来实现和解决无线传感器网络中的数据融合问题。神经网络用于无线传感器网络的数据融合已经被证明是非常有效的。

Reznik 等^[4] 应用多层感知神经网络检测无线传感器网络的信号变化, 判断异常事件的发生。Wilbert 等^[5] 将遗传算法和神经网络算法相结合, 在监测系统的多任务进程管理中取得了较好的表现。模式分类可以在节点中独立判断, 簇首节点搜集其它节点的分类信息, 进一步融合数据, 节省了通讯能量, 实验显示该算法具有很强的数据鲁棒性。Julio Barbancho 等^[6] 将自组织映射网络 (Self-Organizing Map, SOM) 引入无线传感器网络的路由决策中, 并对引入神经网络的效率进行分析。陈斌等^[7] 结合神经网络和证据理论, 建立基于无线传感器网络的管道泄漏诊断模型。俞黎阳等^[8] 将无线传感器网络的分簇层次结构与神经网络的层次结构相结合, 构造一个三层感知器神经网络对无线传感器网络进

行数据融合, 仿真测试的结果显示该模型可以有效地节省传感器节点的能耗、延长网络寿命。但该模型在分析簇首更替与神经网络参数移交时未考虑簇成员改变和节点死亡所产生的影响。W-T Sung^[9] 利用 BP 神经网络对环境监测系统中的多传感器数据进行融合, 该环境监测系统基于无线传感器网络。实验结果表明, 引入 BP 神经网络方法可以大大降低数据特征维数, 提高环境监测系统的数据融合效率。

本文将 BP 神经网络引入到无线传感器网络的数据融合中, 构建一种新的数据融合算法。为方便讨论, 文中假设 N 个传感器节点随机分布在感知区域内, 用 s_i 表示第 i 个节点, 相应的节点集合为 $S = \{s_1, s_2, \dots, s_N\}$ 。并且该传感器网络具有如下性质:

- (1) 节点部署后不再移动, 所有传感器节点都被事先编排唯一的 ID 号。
- (2) 节点能量相同且不能补充。
- (3) 汇聚节点 (即基站) 唯一且部署在感知区域以外的固定位置。
- (4) 汇聚节点拥有持续的能量供给, 能够使用足够大的功率向所有节点直接发送信息, 而节点发射功率有限。
- (5) 节点能够获知其位置信息。

前 4 项假设是无线传感器网络的典型设置。第 5 项假设十分必要, 因为节点通常需要获取它的位置信息, 尤其当传感器网络簇首更替、簇成员发生变化或节点死亡时, 确定节点的位置十分重要。

2 基于神经网络的数据融合

2.1 BPNDA 算法的模型

BPNDA 算法基于无线传感器网络的 LEACH 分簇路由协议, 在整个网络范围中按照一定的规则选举簇首, 形成分簇结构。在这种分簇结构下, 传感器节点采集到的大量原始数据将首先被发送给自己所在簇的簇首节点。BPNDA 算法在簇首节点和成员节点间利用 BP 神经网络算法对采集到的原始数据进行融合处理。

图 1 给出了 BPNDA 算法的模型示意图, 该结构采用了三层 BP 神经网络模型, 对应无线传感器网络中的一个簇。理论已经证明, 三层 BP 神经网络, 只要隐层节点数足够多, 就具有模拟任意复杂的非线性映射的能力^[10]。考虑到簇成员节点具有对采集数据进行滤波等预处理的功能, 文中将 BPNDA 算法的输入层设于簇成员节点中, 而隐层和输出层位于簇首节点中。

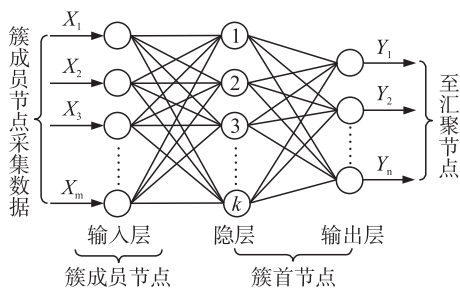


图1 BPNDA数据融合算法的模型图

假设无线传感器网络中一个簇内有 m 个簇成员节点,那么该神经网络模型共有 m 个输入层神经元。输出层神经元的数量 n 可以根据实际应用的需要进行调整,一般由研究对象的输出信息来确定,与簇成员节点的数量没有必然联系。而隐层神经元数量 k 的确定与求解问题的要求、输入输出神经元数量都有直接的关系,是神经网络设计中非常重要的一个环节。但是,到目前为止,BP网络隐层神经元数量的确定尚无成熟的理论指导,本文采用“试算法”来确定隐层神经元的数量:根据经验公式确定隐层神经元数量的范围,设计一个隐层神经元数量可变的BP神经网络结构,基于相同的训练目标精度、训练样本和测试样本数,通过对比不同隐层神经元数量下的训练次数和识别精度,选取最佳的隐层神经元数。最佳的隐层神经元数量由经验公式 $k = \sqrt{m+n} + \beta$ 确定,其中, β 为 $[1, 10]$ 之间的常数。

根据这样一种三层BP神经网络模型,BPNDA数据融合算法首先在每个传感器节点对所有采集到的数据按照输入层神经元函数进行初步处理,然后将处理结果发送给其所在簇的簇首节点。簇首节点再根据隐层神经元函数和输出层神经元函数进行进一步处理。最后,由簇首节点将处理结果发送给汇聚节点。

从无线传感器网络的整体来看,簇成员节点就相当于最底层的神经元,簇首节点则是起到汇聚作用的中间神经元,而整个网络则可以被看成是一个复杂的神经系统。

2.2 簇首选取及分簇

已有分簇算法基本上都是每轮循环都要构造簇,这无疑会增加构造簇的开销,考虑利用网络自身拓扑结构,得到一个稳定的簇结构,无需每轮循环都构造簇,从而减少网络能量消耗。因此,Heinzelman还提出了LEACH-F算法^[11];簇组织完成后即固定不变,而簇首实行轮转,这种方法减小了成簇开销,并且簇首分布总体上较为均匀。

文中BPNDA算法结合LEACH-F算法的思想,

在网络初始布置的时候,通过一次运行LEACH或LEACH-F分簇算法,得到一个拥有稳定簇结构的网络,直到有大量新的节点加入或失效的节点退出之前,都不需要再运行此算法。算法的运行和数据的存储都可以在Sink汇聚节点中解决,这样可以大大减少传感器节点的运算和存储负担。

由于簇首的负担较重,BPNDA算法在构造簇时,特意选择剩余能量较高的节点作为簇首。令 s_i 为任意的一个候选簇首。 s_i 根据自身到汇聚点的距离信息计算它的竞争区域,区域的半径记作 R_c 。下面定义候选簇首之间竞争的规则:

在竞选过程中,若候选簇首 s_i 宣布其竞选获胜,则在 s_i 的竞争半径 R_c 内的所有候选簇首均不能成为最终簇首,需要退出竞选过程。

每个候选簇首维护一个邻簇首集合 S_{CH} ,在这个集合内依据当前各节点的剩余能量高低竞争选出最终簇首。在BPNDA簇首竞选中,候选簇首 s_i 的邻簇首集合 $s_i \cdot S_{CH}$ 包括与 s_i 具有竞争规则所约束的竞争关系的所有候选簇首节点:

$$s_i \cdot S_{CH} = \{s_j | s_j \text{ 是候选簇首, 且 } d(s_i, s_j) < \max(s_i \cdot R_c, s_j \cdot R_c)\}。$$

每个节点均以同样的功率发送广播消息,为了节约能量,这个广播半径设为 R_c^0 即可(这保证了节点能够与邻簇首集合内的所有节点正常通信)。下面阐述竞争选取簇首的步骤:

Step 1 依概率在网络中选出部分节点成为候选簇首,参与竞选。普通节点成为候选簇首的概率为 T ,它是一个预先设置的阈值。未参与竞选的节点进入睡眠状态,直到簇首竞选过程结束。

Step 2 每个竞选节点广播竞选消息 COMPETE_HEAD_MSG,消息的内容为节点的ID、竞争半径 R_c 和当前剩余能量 E_R 。

Step 3 节点根据收到的广播消息构建其邻簇首集合 S_{CH} 。

Step 4 当 S_{CH} 构建完成后,节点做出其是否能担任簇首的决策。节点需要等待其邻簇首集合中所有能量比它大的节点先做出决策,然后才能确定自身是否能担任簇首。若节点的剩余能量相等,则以节点的ID作为次比较依据。

Step 5 一旦 s_i 发现它的剩余能量比其邻簇首集合中的节点的剩余能量都高,则它将赢得竞选,并广播获胜消息 FINAL_HEAD_MSG 以通知它的邻簇首。

Step 6 如果 s_i 收到来自 s_j 的获胜消息,且 s_j 是 $s_i \cdot S_{CH}$ 中的一个节点,则 s_i 立即退出竞选,并广播消息 QUIT_ELECTION_MSG 通知它的邻簇首。

Step 7 如果 s_i 收到来自 s_j 的退出消息, 且 s_j 是 s_i S_{CH} 中的一个节点, 则 s_i 将 s_j 从其邻簇首集合中删除。

在簇首竞选过程结束后, 之前未参与竞选的节点从睡眠状态唤醒, 接着竞选产生的簇首向全网广播其竞选获胜的消息 CH_ADV_MSG。普通节点选择簇内通信代价最小亦即接收信号强度最大的簇首, 发送加入消息 JOIN_CLUSTER_MSG 通知该簇首。这样, 网络就形成了分簇结构, 接下来进入簇内部的数据传输阶段。

2.3 神经网络训练及数据融合

BPNDA 算法中的神经元在运行时需要一些运行参数(如权值), 因此, 在完成簇首选取和分簇后, 无线传感器网络进入自治工作状态前, 需对 BP 神经网络进行训练。通常, 神经网络可以通过网络本身学习、训练、调整这些参数。但在无线传感器网络中, 由于受到能量、处理能力和存储空间等资源的限制, 无法在无线传感器网络内部对神经网络进行训练。因此, BPNDA 算法选择在无线传感器网络汇聚节点完成对神经网络的训练。网络训练及数据融合基本步骤如下:

Step 1 簇首节点向汇聚节点传送其簇内节点信息表;

Step 2 汇聚节点依据簇首及其簇内节点信息, 构造 BP 神经网络;

Step 3 BP 神经网络检索样本数据库, 搜集与簇成员信息匹配的样本进行训练, 从而生成相应簇的神经网络参数;

Step 4 汇聚节点将 BPNDA 各神经元参数发送给对应节点, 包括簇首节点和成员节点;

Step 5 分簇稳定工作, 簇首对接收信息进行融合并向汇聚节点传递融合后数据。无线传感器网络运行过程中产生的历史记录也可添加至样本数据库。

图 2 为 BPNDA 数据融合算法的结构图, 其中 w_1, w_2 为权值矩阵; b_1, b_2 为阈值矩阵; f_1, f_2 为激励函数, 其中 f_1 为 tansig 型函数, f_2 为 purelin 型函数; x_1, x_2 为各层神经网络输入, $x_2, output$ 为各层神经网络输出。对于多层神经网络, 上一层的输出为下一层的输入。 $x_2 = f_1(w_1 \times x_1 + b_1)$, $output =$

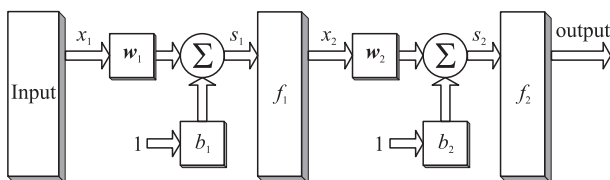


图 2 BPNDA 数据融合算法的结构图

$f_2(w_2 \times x_2 + b_2)$, output 为代表感知区域观测对象的特征值。

当 BP 神经网络经离线训练好后(由汇聚节点完成), 神经网络各层的权重、阈值便确定下来, 再由汇聚节点将确定的权值及阈值发送至对应节点。无线传感器网络分簇即可利用训练生成的神经网络参数进行在线计算, 实现数据融合。由神经网络的输入矢量 Input(即簇成员的采集数据)求解输出矢量 output(即数据融合结果)的步骤如下: ① $s_1 = \sum x_1 \times w_1 + b_1$, 其中: $x_1 = \text{Input}$; ② $x_2 = f_1(s_1) = \frac{2}{1 + e^{-2s_1}} - 1$; ③ $s_2 = \sum x_2 \times w_2 + b_2$; ④ $output = f_2(s_2) = s_2$ 。

文献[12]表明, 每传输 1 byte 数据所消耗的能量可以用来执行数千条 CPU 条指令, 而且所消耗的时间也少得多。因此, 虽然 BPNDA 数据融合算法增加了有限的计算量, 但其大幅降低了数据通信量, 将有效提高无线传感器网络的工作周期。

3 算法的仿真测试

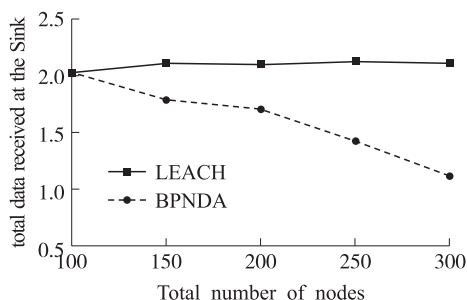
本文选用 NS-2 仿真工具对 BPNDA 数据融合算法进行仿真测试, 以液压系统的状态监测为应用实例, 布置在液压系统中不同监测位置的流量、压力和温度传感器节点不间断采集其相对应信号, 而后经 BPNDA 进行融合处理后由簇首节点向汇聚节点发送液压系统的工作状态。在神经网络训练过程中, 当液压系统正常时, 设定神经网络输出为 0, 否则为 1^[13]。因此, 输出层只设定 1 个神经元, 用来输出液压系统的状态。

BPNDA 算法的性能主要通过 Sink 节点接收的数据量与平均节点能耗两个方面进行仿真评估。为了测试不同网络规模下 BPNDA 的性能, 选取 5 种具有不同的节点密度、不同规模的无线传感器网络场景: 100、150、200、250 和 300 个传感器节点随机分布在一个 100 m×100 m 大小的区域内, 具有较好的可比性。为最大限度地减少网络随机部署的影响, 在每种测试方案中, 都生成 10 个不同的网络拓扑结构, 最后取其运行结果的平均值, 并将其与 LEACH 算法进行性能对比。节点的具体的仿真参数如表 1 所示。由于在 BPNDA 算法中, 簇首节点需发送的数据仅为代表液压系统工作状态的 0 或 1, 不同于簇成员节点发送的流量、压力和温度数据; 同时, 在传感器网络中, 数据字段的长度通常为 2 byte, 一个数据字段完全可以表示出液压系统的工作状态, 所以在仿真时, 将簇首节点发送数据包的长度设为 2 byte。

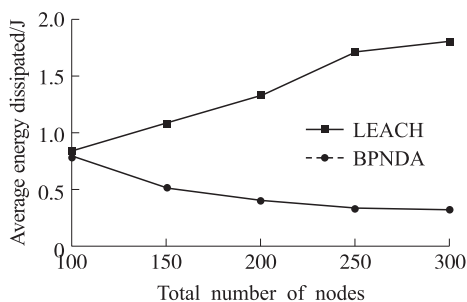
表1 仿真参数表

参数	取值
网络范围	100 m×100 m
Sink 节点坐标	(50,175)
初始能量	2 J
数据包长度	500 byte
分簇消息长度	25 byte
包头长度	25 byte
E_{elec}	50 nJ/bit
ϵ_{fs}	10 pJ/(bit·m ²)
ϵ_{mp}	0.001 3 pJ/(bit·m ⁴)
E_{DA}	5 nJ/(bit·signal)
仿真时间	200 s

图3反映了BPNDA算法对网络性能的影响。其中图3(a)给出Sink节点接收到数据量与不同网络规模的关系,由图可见,随着网络规模的增大,在LEACH算法中,由于信道和通信速率的限制,大量的冲突数据包被丢弃,所以Sink节点接收的数据量在增加到一定程度后保持基本稳定;而采用BPNDA数据融合算法,簇首节点发送的数据包(25 byte + 2 byte = 27 byte)远小于簇成员节点以及LEACH算法中簇首节点发送的数据包(25 byte + 500 byte = 525 byte),因此可以有效减少数据通信量,并且随着传感器节点数量的增长,其融合效果将得到更好的体现,Sink节点接收的数据量也随之下降。



(a)BPNDA对Sink接收数据量的影响



(b)BPNDA对节点平均能耗的影响

图3 BPNDA对网络性能的影响

图3(b)反映了不同网络规模下,传感器节点平均能耗的对比结果。通过对该结果的分析,可以得出,采用BPNDA数据融合算法,传感器节点平均能耗明显低于使用LEACH算法时的平均能耗。并且,随着传感器节点数量的增加,使用BPNDA算法时节点平均能耗随数据通信量的下降而下降,而使用LEACH算法时节点平均能耗则随节点数量的增加而上升,两者之间的差值逐渐增大。

4 结论

对于无线传感器网络,基于分簇的网络结构是一种较好的选择,在簇内可以结合数据融合功能来提高网络性能。本文提出了一种适用于分簇无线传感网络的数据融合算法BPNDA,仿真测试表明,该算法能够有效减少网络中的数据通信量,从而降低传感器节点的能耗,延长网络的生命周期。

参考文献:

- [1] Intanagonwivat C, Govindan R, Estrin D. Directed Diffusion: a Scalable and Robust Communication Paradigm for Sensor Networks [C]//New York: MobiCom'00, 2000:56-67.
- [2] Andr L L de Aquino, Carlos M S Figueiredo, Eduardo F Nakamura, et al. Data Stream Based Algorithms for Wireless Sensor Networks Applications [C]//Ontario: 21st International Conference on Advanced Networking and Applications (AINA'07), 2007:869-876.
- [3] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-Efficient Communication Protocols for Wireless Microsensor Networks [C]//Proceedings of 33rd Hawaii International Conference on Systems Science, Washington, DC, 2000:8020-8030.
- [4] Reznik L, Von Pless G, Al Karim T. Intelligent Protocols Based on Sensor Signal Change Detection [C]//Proceedings of Systems Communications, 2005:443-448.
- [5] van Norden W, de Jong J, Bolderheij F, et al. Intelligent task Scheduling in Sensor Networks [C]//Proceedings of 8th International Conference on Information Fusion, 2005.
- [6] Julio Barbancho, Carlos León, F J Molina, et al. Using Artificial Intelligence in Routing Schemes for Wireless Networks [J]. Computer Communications, 2007, (30):2802-2811.
- [7] 陈斌, 万江文, 吴银锋, 等. 神经网络和证据理论融合的管道泄漏诊断方法 [J]. 北京邮电大学学报, 2009, 32(2):5-9.
- [8] 俞黎阳, 王能, 张卫. 无线传感器网络中基于神经网络的数据融合模型 [J]. 计算机科学, 2008, 35(12):43-47.
- [9] Wen-Tsai Sung. Employed BPN to Multi-Sensors Data Fusion for Environment Monitoring Services [J]. Autonomic and Trusted Computing, 2009, (6):149-163.
- [10] Omid O, Judith D. Neural Networks and Pattern Recognition [M]. San Diego: Academic Press, 1998.
- [11] Heinzelman W. Application-Specific Protocol Architectures for Wireless Networks [D]. MIT, 2000.
- [12] Min R, Bhardwaj M, Seong-hwan Cho, et al. Energy-Centric

Enabling Technologies for Wireless Sensor Networks [J]. IEEE Wireless Communications, 2002, 9(4): 28-39.

[13] 黄志坚,袁周. 液压设备故障诊断与检测实用技术[M]. 北京:

机械工业出版社, 2005, 7: 215-217.

[14] 李成法,陈贵海,叶懋,等. 一种基于非均匀分簇的无线传感器网络路由协议[J]. 计算机学报, 2007, 30(1): 27-36.



孙凌逸(1979-),男,第二炮兵工程学院兵器发射理论与技术专业博士研究生,主要研究方向为兵器发射理论与技术及无线传感器网络, sunlyi79@163.com;



黄先祥(1940-),男,第二炮兵工程学院教授,博士生导师,中国工程院院士,主要研究方向为兵器发射理论与技术、发射系统仿真与自动检测技术;



蔡伟(1974-),男,第二炮兵工程学院,副教授,主要研究方向为智能检测、系统仿真及无线传感器网络;



夏梅尼(1981-),女,中国空间技术研究院西安分院,工程师,硕士,主要研究方向为通信工程。