

订阅DeepL Pro以编辑此演示文稿。  
访问[www.DeepL.com/pro](https://www.deepl.com/pro?cta=edit-document)，了解更多信息。

灾后搜救行动中基于强化学习的无人机基站轨迹优化研究

Shiye Zhao，*学生会员，IEEE；*Kaoru Ota，*会员，IEEE；*Mianxiong Dong，*会员，IEEE*

***摘要--由于*灾难的发生，地面基站（TBS）会有部分崩溃。一些用户设备（UE）将得不到服务。部署无人机（UAV）作为空中基站是一种快速覆盖用户设备的方法。但现有的方法只提到了无人机的覆盖。在这些情况下，它们侧重于在所有TBS都不再工作的灾后地区部署无人机。关于可用的TBS和无人机的结合的研究很有限。我们提出了将无人机与可用的TBS合作部署为空中基站的方法。并通过强化学习提高覆盖率。此外，在实验中，我们首先用平衡的迭代减少和使用层次的聚类（BIRCH）对UE进行聚类。最后，通过Q-learning实现基站对UE的更好覆盖。**

***关键词：无人机*基站，Q-learning，无人机的部署和移动。**

1. 简介

arXiv:2202.10338v1 [cs.LG] 2022年2月17日

**W**

IDE-规模的自然灾害，如洪水、海啸和地震，发生在不同的地方，并产生破坏性的后果。在突如其来的灾难中，受害者往往需要尽快与外界联系以获得援助。然而，就像飓风、洪水或地震一样，地面通信网络可能部分或完全中断[[1](#_bookmark23)]。而余震的存在是一个潜在的威胁，导致现有的无线通信网络不能被完全依赖[[2](#_bookmark24)]。应急通信应及时部署。在灾后的情况下，地面交通通常被阻断。紧急通信车辆无法到达现场提供紧急通信服务。基于地面的通信方法对于维持灾后地区人们的可靠通信服务是具有挑战性的。由于复杂的灾后情况，该方法应该被灵活地实施[[3](#_bookmark25)]。换句话说，我们需要可靠和灵活的无线通信，在任何复杂的灾后条件下充当应急网络，用于灾后救援和应急通信[[4](#_bookmark26)]。

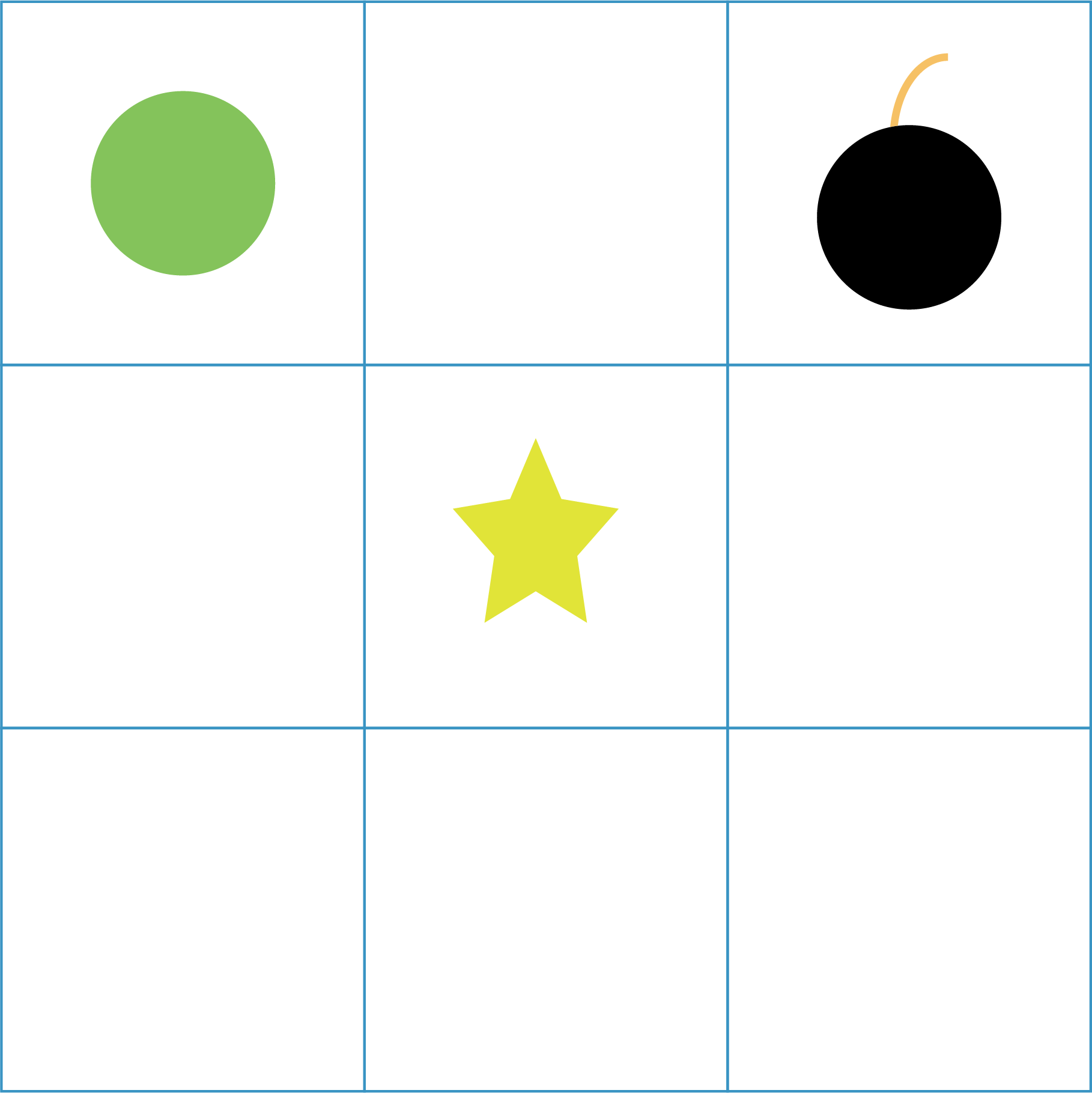
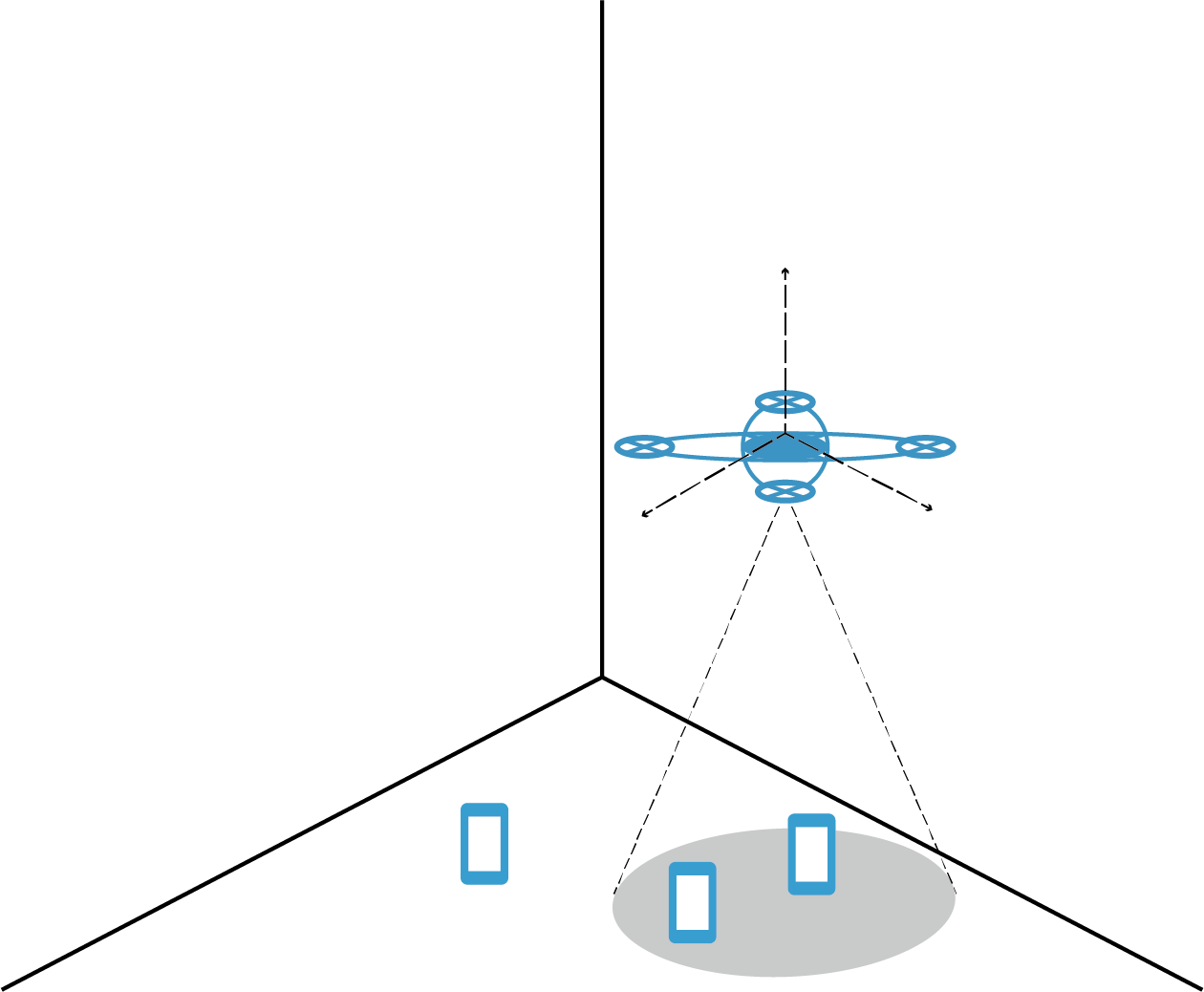
部署无人机作为空中基站可以解决这个问题。

无人机主要用于军事任务，但随着电子技术的发展，无人机因其灵活性而在侦察方面得到迅速发展[[5](#_bookmark27)]。随后，信息技术的发展拓宽了无人机的应用范围，因为它们可以代替人类完成空中作业，如物流、农业、救援、测绘等[[6](#_bookmark28)]。同时，更大的带宽、更低的延迟和更高密度的连接移动网络也为无人机提供了更多的机会。例如，无人机可以作为空中基站，向所需区域的地面用户设备（UE）提供无线通信[[7](#_bookmark29)]。由于无人机的敏捷性，无人机基地

Shiye Zhao, Kaoru Ota和Mianxiong Dong是日本室兰工业大学科学和信息学系的，（电子邮件。20043038@mmm.muroran-it.ac.jp; ota@csse.muroran-it.ac.jp; mx.dong@csse.muroran-it.ac.jp）。

站可以可靠而有效地以低成本提供灵活的通信网络[[8](#_bookmark30)]。然而，由于无人机电池的限制和灾后场景的时间敏感性，需要准确和迅速地部署无人机。放置优化的问题应该得到解决[[9](#_bookmark31)]。换句话说，如何部署无人机以在更短的时间内覆盖更多的地方是我们关注的重点。此外，部署时应考虑TBSs部分故障的情况，以减少无人机的额外成本，因为有些TBSs在灾后仍能工作。换句话说，如何在更短的时间内部署无人机来覆盖更多的用户，并与现有的TBS合作是我们关注的重点。

我们提出了一种方法来部署无人机作为空中基站，与这些现有的TBS合作，基于Q-learning为灾后地区的所有UE服务。在灾后的情况下，一些TBS不会被撞毁。这些现有的未崩溃的TBS仍然可以继续工作，提供无线通信服务。因此，我们可以在部署无人机之前确定UE的连接状态，对这些没有服务的UE进行分类。然后在受损的TBS区域为UE部署无人机，与现有的TBS合作，建立一个应急通信系统，为所有UE服务。在Q-learning的典型应用中，有一个像图[1](#_bookmark0)中（a）的网格世界游戏，其中绿色的圆圈是玩家，玩家遇到黄色的星星就会赢得游戏；相反，玩家遇到黑色的炸弹就会失去游戏。无人机的部署和移动可以被视为这样一个三维网格世界的游戏，如图[1](#_bookmark0)中的（b）。无人机作为一个玩家，为UE提供紧急通信服务。当所有UE在规定时间内得到服务时，玩家将赢得游戏，而当所有UE在规定时间内未被覆盖时，玩家将失去游戏。同时，为了加快Q-learning的步骤，我们在Q-learning之前实现了带有平衡迭代还原和使用层次结构的聚类算法（BIRCH）。聚类后，对于每个无人机来说，需要无线通信服务的UE数量减少，那么无人机的工作范围就会缩小，促进无人机在Q-learning中找到覆盖UE的优化位置。

1. 网格世界的环境 （b）无人机的环境 图1.网格世界和无人机之间的环境

学习。以下是本文的贡献。

* 考虑到TBS和空中基站的合作。关于无人机与地面基站共存的部署的文献很少。现有的TBSs仍将为UE提供无线通信服务。考虑到现有的TBSs，无人机的部署将减少成本并提高效率。
* 通过BIRCH的预处理加速Q-learning的训练。由于BIRCH的实施，原来未被碰撞的TBS所覆盖的区域将根据空中基站的覆盖范围被划分为更小的区域。无人机将更有效地找到它们的最佳位置，加速了Q-learning的训练。
* 动态地部署无人机。无人机的部署是基于BIRCH的聚类结果，而BIRCH不需要预先输入聚类的数量。因此，分配的无人机数量将随着任务的要求而变化。如果空中基站的覆盖范围有限或地图的规模较大，就会分配更多的无人机。否则，将部署较少的无人机。

本文的其余部分组织如下。在第二节，我们介绍了无人机部署的相关工作。在第三节中，我们提出了无人机与现有TBS合作以覆盖所有UE的模型。第四节，研究了无人机在蜂窝网络中部署的聚类和强化学习。第五节介绍了无人机的部署和移动的模拟结果。最后，在第六节中提出了结论。

1. 相关作品

对于行业来说，亚马逊的Prime Air[[10](#_bookmark32)]和Alphabet的Project Wing[[11](#_bookmark33)]是在送货服务中部署蜂窝连接无人机的杰出案例。EE和诺基亚使用无人机携带一个小型基站来提供无线通信[[12](#_bookmark34)]。谷歌的Loon项目[[13](#_bookmark35)]测试了将无人机用于无线通信；同时，Qual-comm和AT&T合作在第五代无线网络中部署了无人机[[14](#_bookmark36)]。关于无人机基站的文献已经被广泛开发。在[[15](#_bookmark37)]中，作者结合建筑几何学和刀锋衍射理论，提出了一个A2G LoS模型。[[16](#_bookmark38)]提出了一种优化无人机高度的方法，以实现最大的覆盖，当

无人机作为一个空中基站。而[[17](#_bookmark39)]将[[16](#_bookmark38)]从单个无人机扩展到两个无人机。关于优化无人机基站的部署也有大量的文献。例如，[[18](#_bookmark40)]讨论了优化具有最大用户的无人机的三维位置。作者在[[19](#_bookmark41)]中提出了一种启发式算法，以最小的数量覆盖所有用户来部署无人机。在文献[[9](#_bookmark31)]中，作者提出了一种按顺序部署无人机的算法，从区域的边缘开始，向内螺旋式上升，直到覆盖所有的UE。[[20](#_bookmark42)]利用Koebe-Andreev-Thurston定理，将无人机部署在其覆盖区域最大的位置。[[21](#_bookmark43)]研究了以用户为中心和以网络为中心的无线回程，并为其找到了最佳的回程软件布置。

在[[22](#_bookmark44)]中，作者提出了推进能量消耗与飞行速度、方向和加速度的理论模型。在[[23](#_bookmark45)]中，将无人机的部署问题分为垂直和水平两个维度，以最小的功率为最大的UE提供无线通信。[[24](#_bookmark46)]在没有用户分布的先验信息的情况下，使用扫荡和搜索用户集群的概念来解决部署有限数量的无人机以覆盖尽可能多的UE。在[[25](#_bookmark47)]中，提出了将强化学习用于无人机部署的方法，通过观察无人机的三维坐标获得最大的体验质量（QoE）。[[26](#_bookmark48)]使用特定的无人机与Q-learning来处理无人机由于电池有限而导致的任务连续性限制。

然而，上述的这些方法仅仅关注于无人机作为空中基站，而TBS完全被破坏的情况。正如我们之前提到的，在灾难发生后，一些TBS可能仍然正常工作。这些可用的TBS仍将为UE提供服务，而不考虑未中断的TBS的存在的无人机部署将导致额外的成本。总之，我们应该根据灾后地区的情况部署无人机作为空中基站，与现有的TBS合作，覆盖灾后地区的所有UE。

1. 系统模型

灾后应急通信的无人机部署情况如图[2](#_bookmark2)所示。灾害发生前，UE由TBS提供服务，如TBS1 。一旦TBS因灾崩溃，如TBS2 ，他们就无法继续为图[2](#_bookmark2)中白圈区域的UE提供服务。同时，由TBS1 服务的UE如果能够继续工作，仍然会被连接，我们不需要为他们部署无人机。由于无人机的覆盖范围有限，不可能为未工作的TBS部署无人机来覆盖该区域。因此，我们需要对未连接的UE进行集群，并将TBS覆盖的原始区域2 ，划分为图[2](#_bookmark2)中的区域1 ，区域2 ，区域3 和区域4 。最后，在这些划分的区域内部署无人机，并调整无人机的位置以获得更好的性能，从而连接所有UE。

为了对这些未服务的UE进行聚类，我们需要区分UE的状态是否需要无线应急网络，如*csbm* 0*,* 1 *,* 0 *< b < N* + *J,* 0 *< m < M* 其中b和m是*bth* 基站的索引和

*∈ { }*

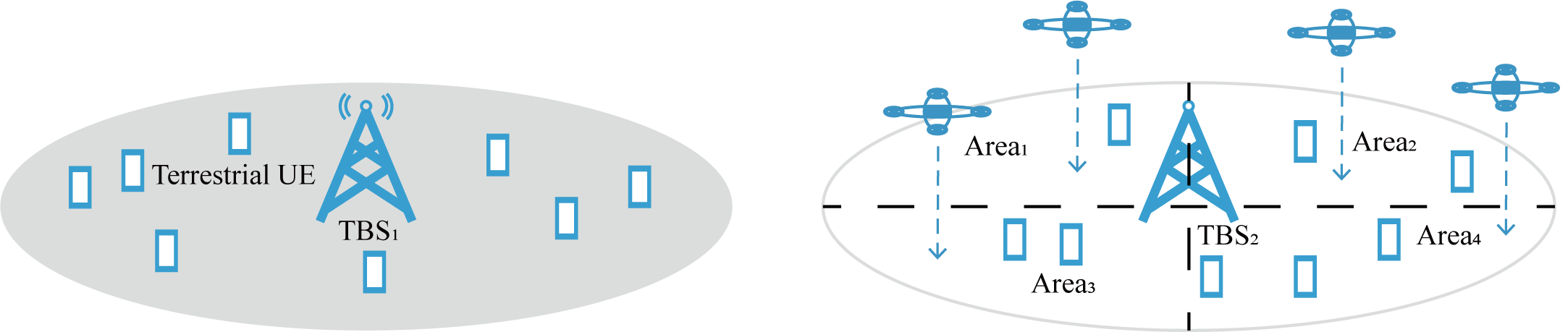
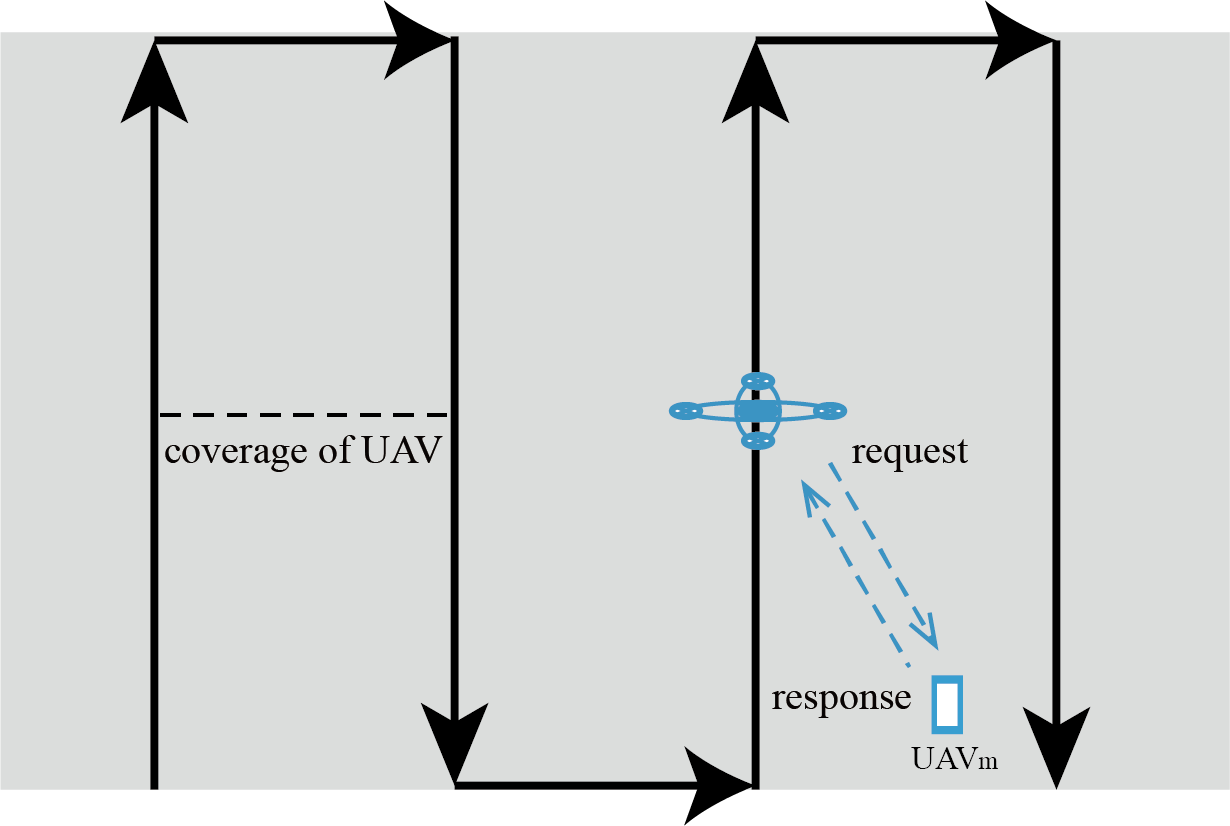


图2.灾后情景

*mth* UE,分别，M是UE的数量。同时，*b csbm*  0*,* 1 .UE*m* 将只选择一个基站的服务。当发射器*b* 和UE*m* 之间没有无线连接时，*csbm* = 0，否则，*csbm* = 1。那么

Σ *∈ { }*

有UE*m* 和TBS的连接状态。

*b=N* +J

*csm* =

*b=ΣN* +j

*csbm*  (1)

当UE*m* 未与TBS连接时，*csm* = 0；如果UE*m* 与TBS连接，*csm* = 1。正如我们前面提到的，一些TBS在灾后仍然可以为UE服务，对于这些UE来说，*csm* = 1。无人机可以被部署来为那些*csm* = 0的UE提供服务。此外，当我们对UE进行集群时，我们需要UE的位置。这里应该实施一种机制来扫描环境并获得UE的状态和信息。为了覆盖整个地图，我们部署无人机以图[3](#_bookmark3)所示的 "之 "字形扫描区域，间隔时间以无人机的覆盖范围为准。在人字形扫描过程中，无人机将广播请求信息。然后，*csm* = 0的UE将对请求信息作出回应。响应信息将包含UE的信息 *U* =U 1 *, U*2 *, ..., Um , ..., UM , Um* =loc *m , csm* 其中*locm* 是UE*m* 的位置，而中心会

*{ } { }*

图3.人字形清扫的机制

*uav*1 *, ..., uavn , ..., uavN , tbs*1 *, ..., tbsj , ..., tbsJ*

*{ }*

表示基站，包括无人机和TBS，其中*N*和*J*分别为无人机和TBS的数量。每个UE到基站的距离被表述为

根据他们的信息进行聚类，即*聚类*=

*{clus*1 *, clus*2 *, ..., clusn , ..., clusN }, clus ∩clusij* = ∅ if *i /= j*。

*b*

*dbm*

= q(*xb*

* *xm*

)2 + (*yb*

* *ym*

)2 + *z*2 *。* (2)

集群后，无人机会调整位置，为UE提供更好的服务。无人机辅助蜂窝网络的传输情况如图[4](#_bookmark4)所示，在图中我们部署了多架无人机作为空中基站与现有的TBS合作，以覆盖灾后地区的所有受害者。对于图[4](#_bookmark4)中的地面UE来说，接收的信号包括两部分，从空中基站到地面UE（A2G）的信号，在图[4](#_bookmark4)中分别表示为A2G1 和 A2G2 。

和TBS到地面UE（G2G）的通道，表示为G2G

其中*x*、*y*和*z*是三维共线的值，*b*表示*bth* 基站的指数。平均而言，我们制定*Dum* = *d*1*m , d*2*m , ..., dnm , ..., dNm* 和 *Dmm* = *d*(*N*+1)*m , ..., d*(*j*)*m , ..., d*(*N*+*J*)*m*  来表示距离无人机和TBS的距离。

对于 *csm* = 1 的 UE*m* ，连接是基于 TBS 和 UE*m* 之间的距离。我们重新定义TBS和UE之间的距离为

*{ }*

*{ }*

在图[4](#_bookmark4)中。在收到来自不同基地的信号后

*dm* =(*d*(*N +*j)*m,cs*(*N* +j)*m* = 1& *d*(*N* +j)*m ≤dthrbs,*

(3)

用户选择与基站建立无线通信链接，以提供最佳服务。换句话说，如果A2G1 对地面UE表现最好，那么A2G2 和G2G对地面UE都是干扰信号。同样地，如果A2G2 表现最好，A2G1 和 G2G 是干扰，如果 G2G 表现最好，A2G 对地面 UE 是干扰。

*jm*

*∞，其他。*

为了计算方便，当用户离开TBS的覆盖范围时，长度被设置为无限。在选择基站时，UE*m* 会考虑距离*dmjm* 满足的TBS。

为了获得更好的服务，UE会比较基站的性能。距离较近的无人机将提供更好的服务，因为损失较少，而且更强大

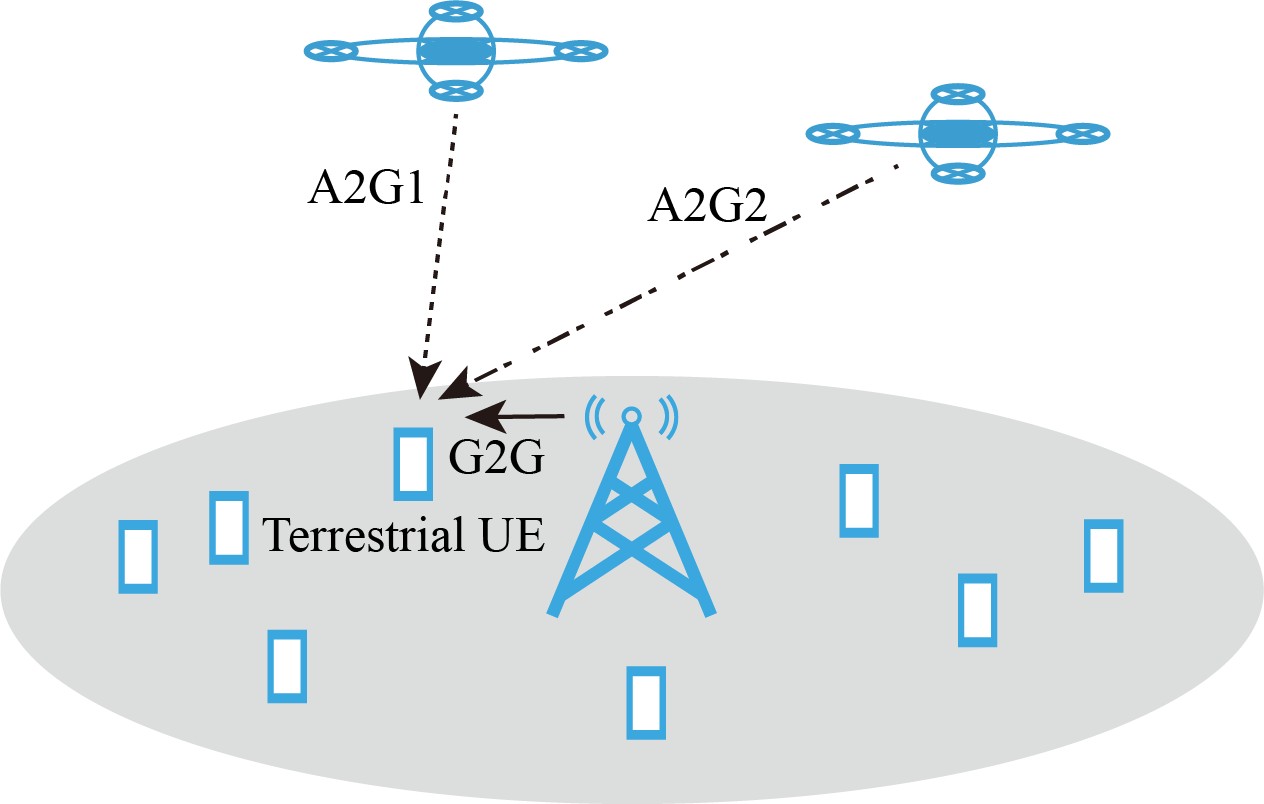
*dmjm* = min *dmjm , dmjm*

*j*

*∞* (4)

信号。我们首先需要的是计算UE到基站的距离。 我们制定*BS* =

考虑到G2G的传播会与反射结合在一起，我们选择了《双射线地面模型》。

无人机在A2G信道中传输的信号包含两部分：视线传播（LOS）和非视线传播（NLOS）。UE会收到包含上述两部分的信号。而LOS和NLOS传播的传输分别发生在不同的可能性上。A2G信道中LOS链路的可能性被表述为

*Plos*(θnm) =

1

1 + *α* exp*(-β*[θ*nm - α*])

*,* (9)

图4.A2G和G2G模型的情景

在我们的模型中，无线通信代表G2G信道的传播。

*|Ejm |2*

其中*α*和β是由环境决定的常数，如乡村或城市，此外θ*nm* 是无人机和UE之间的角度，由θ*nm* = arcsin(*yn /dunm* )计算得出。然后，我们可以从以下方面计算出NLOS链接的可能性

*Pnlos (θnm* ) = 1 *- Plos (θnm* )*。* (10)

在传播模型中，由于多个反射信号造成的多径阴影比其他成分要小，我们只关注LOS、NLOS的路径损耗。

*Prbjm* = 10 log(

*Arj* ）*。* (5)

*η*

和自由空间的传播。传播模型中空中基站与地面UE之间的路径损耗可以是

*η*是一个常数，表示自由空间的阻抗，和

*Arj* 是接收天线的有效表面，可以

拟定的

通过以下方式计算

*λ2*

*Arj* = *tbsj Gr*

4π

，其中

*λtbs*

的波长。

*Lnm* = 20 log(

4πf*c dunm*

*c*

)+ *Plosµlos* + *Pnlosµnlos。*

(11)

来自地面基站的信号，而*Gr* 是来自地面用户的接收天线的增益。此外，*Ej m*是电场强度，公式为

4 *πhj hm √30Pt Gt*

*E* = *,* (6)

*λjs d'*2

*jm*

其中*Gt* 和*Pt* 分别为发射天线的增益和发射功率，*hj* 和*hm* 为发射器和接收器的高度，*d′ jm*为它们之间的距离。

由于灾后由正常工作的TBS服务的UE在计算选择空中基站的最小距离时不会重新选择无人机进行连接，因此我们只考虑未连接的UE选择基站时的*Dum* 。

对于*csm* = 0的UE*m* ，UE的初始连接状态表示地面基站的受损情况，因为用户在灾害发生前一直在连接地面基站，并且*csm* = 1，一旦基站崩溃*csm* = 0。因此，空中基站将为UE建立紧急无线通信网络，*csm* = 0。每次无人机移动时，如果无人机与无人机所在区域的UE之间的距离低于无人机的最大范围dthruav，则为

其中*c*是光速，*fc* 是信号的载波频率，此外*μLOS* 和*μNLOS* 是代表LOS和NLOS中的衰减的常数值。有了我们上面计算的空中基站到地面用户的路径损耗，它们之间的功率也可以通过以下方式来计算

*Prnm* = 10 log *Ptn - Lnm 。* (12)

其中*Ptn* 是空中基站的发射功率。而*Prnm* 是地面用户的接收功率。

如图[4](#_bookmark4)所示，地面上的UE会一直接收来自基站的信号，如A2G1 ，A2G2 ，和G2G。我们假设基站的选择方法是基于上述的距离。UE*m* 的干扰如图[5](#_bookmark5)所示。

在图[5](#_bookmark5)中，圆圈的内、外半径dthruav和d*thrbs* 是无人机和TBS的最大连接距离。对于UE1 ，它处于有源TBS1 的覆盖范围内；因此，来自无人机的A2G1 和A2G2 是干扰。同时，UE2 不在 TBS2 的覆盖范围内，但被 UAV3 和 UAV4 覆盖。假设 UE2 在 UAV4 *clus*4 的集群中，G2G2 和 A2G3 对 UE2 都是干扰。有效信号的功率为 *Pe* 1 = *Prb*11 和 *Pe* 2 = *Pr*42 ；同时，干扰信号的功率为

*cs* =(1*,Um ∈clusn* & *dunm ≤dthruav。*

(7)

是 *Pi* 1 = *Pr*11 + *Pr*21 和 *Pi* 2 = *Prb*22 + *Pr*32 。考虑到

*纳米*

0*，其他。*

同样地，我们也将无人机和UE之间的距离表述为

我们在上文中所提出的，连接的力量和

可以得到干扰。然后，信号与干扰加噪声的比率（SINR）可以通过以下方式计算出来

*PemSINR* = *,* (13)

(*dnm, dnm ≤dthruav ,*

*m 貔貅*

+ *Pnoise*

*dunm* =

*m*

(8)

*∞，其他。*

其中*P*

*噪声*

是噪声的干扰功率。

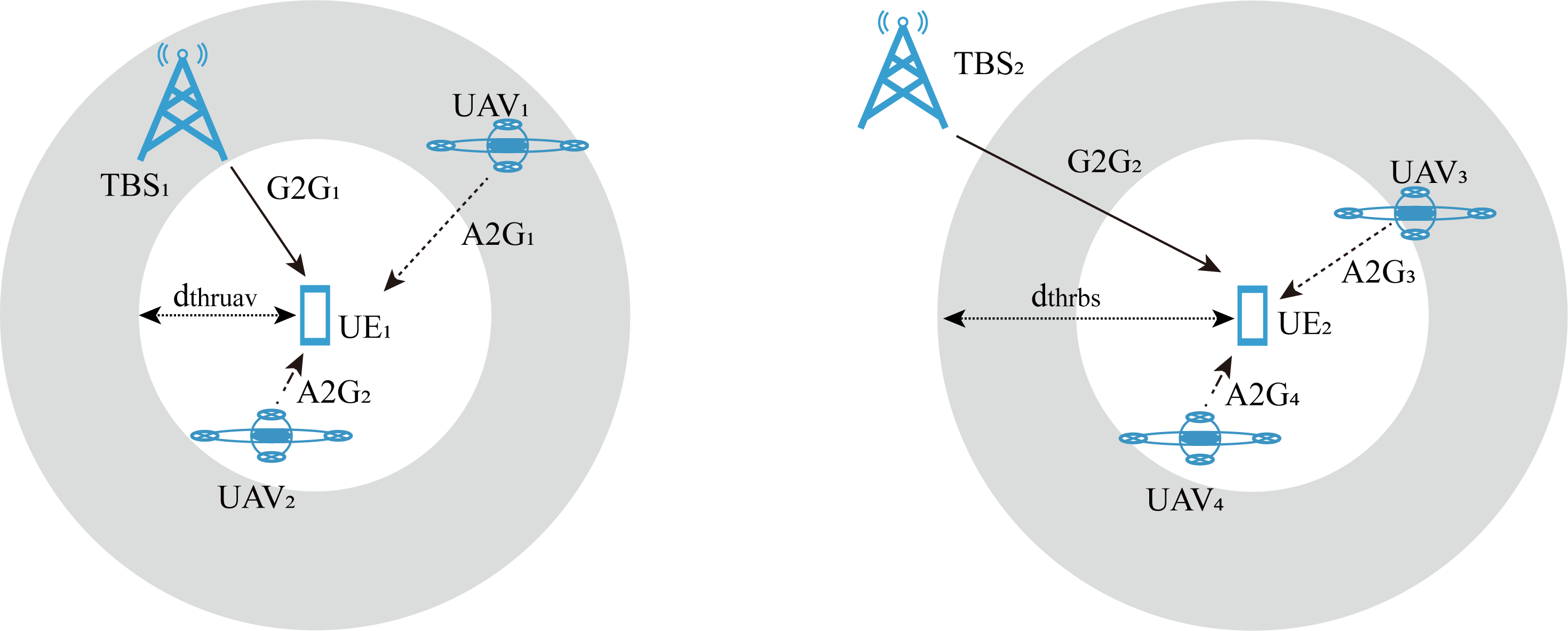


图5.最小距离和阈值的干扰



图6.方法的步骤

同时，用户的比率可以通过Shan-non的公式来计算，其公式为

*Bb*

的无人机，以获得我们在[第三节](#_bookmark1)谈到的最大数量的UE。

在部署无人机之前，我们需要首先检测需要未连接的UE的位置的区域。有了未连接UE的位置，我们用聚类算法对其进行聚类，如图[7](#_bookmark8)。在灾难发生之前，所有的UE都像往常一样由TBS提供服务。由于灾害的影响，TBS1 不能继续工作，由TBS1 服务的UE都没有连接。为了服务灾后地区的所有UE，我们需要建立一个可用的应急无线通信

*率bm* =

*M*

*b*

log2 (1 + *SINRm* )*。* (14)

网络。对于连接的UE，如UE2 , UE3 , UE4 ，它们是

*Bb* 是发射器的带宽，*Mb* 是发射器的接收机数量*b* 。对于UE*m* ，最终速率也可表示为

Σ

*率m* = *rbm csbm 。* (15)

*b*

我们所要解决的是最大限度地提高*Mn* ，并有更高的

*率m* ，因为

max(Σ *Mn* + Σ *Mj* )*。*

, *∈*

仍然与一个TBS连接，如图[7](#_bookmark8)中的TBS2 , TBS3 , TBS4 。我们需要部署空中基站来覆盖未连接的UE，如图[7](#_bookmark8)中的UE1 ，但那些连接的UE。在[第三节](#_bookmark1)中，我们将UE的信息表述为*Um* = *locm , csm* ，然后部署无人机进行之字形扫频以获得所有UE的信息。通过我们得到的*Um* ，我们可以直接拾取这些没有服务的UE，其*csm* = 0，并且

完成分类的任务。

*{ }*

由于无人机的覆盖范围比TBS小，我们需要对未连接的UE进行聚类，并划分区域

*s.t.*

*n*

*j*

*sn Sn*

*s* = *s 如果 n n′*

*/ 'n n*

, *tn < tth*

(16)

由TBS1 覆盖。在图[7](#_bookmark8)中，UE被聚成两个簇：*clus*1 和 *clus*2 ，其中簇中UE周围的圆圈是无人机的移动范围。为了动态地部署无人机，我们用平衡迭代缩减法对UE进行聚类。

它表示我们将实现大多数受害者可以在所需的时间*tth* ；同时，*无人机n* 的位置在其分配的区域*Sn* ，对于不同的无人机，其位置*sn* 和*sn '* 应该是不一样的。

1. 建议的算法

为了在灾后覆盖更多的用户，我们提出了在TBS部分瘫痪的情况下部署空中基站的方法，包含三个步骤，如图[6](#_bookmark6)所示。如前所述，我们不需要为这些被可用的TBS覆盖的UE部署无人机，所以我们应该在一开始就对哪些TBS不工作进行分类，然后在其他步骤之前挑选这些没有服务的UE。然后，考虑到空中基站的覆盖范围，我们设计适当的无人机分配。最后，优化航迹

和使用层次结构的聚类（BIRCH）。BIRCH基于聚类特征（CF）树对数据进行聚类，该树由具有CF的节点组合而成。CF是一个线性三元数组，包含样本数量、每个样本的每个特征维度的总和向量以及它们的平方之和，其中父节点的CF等于子节点的所有CF之和。BIRCH的过程是用所有数据建立一个CF树。当添加新数据时，在叶子节点中搜索最接近的CF，如果数据在半径阈值下的CF覆盖范围内，则将数据添加到最接近的CF叶子节点中，否则数据将被添加到叶子节点的新CF中。在向CF树中添加数据时，如果非叶子节点的CF数量超过默认的非叶子节点的最大CF数量，则将原节点分为两个节点，选择原节点中距离最远的两个CF作为新节点中的第一个CF，然后再添加其他节点

基于两个叶子的距离。将数据添加到新节点后，更新CF树。同样，如果叶子节点的CF数量超过了叶子节点的最大CF数量，就把它分成两个新的节点并为它们分配CF。然后，我们可以得到每个叶子节点的聚类。聚类后，根据每个聚类的最小包围圈来划分区域，以减少无人机的面积

工作，如图[7](#_bookmark8)中的*clus*1 和*clus*2 。

**算法1** 无人机部署的预先评估

**要求：**UE的位置和TBS的状态

根据*UE的*zigzag存储信息，移动无人机以覆盖灾后地区*m* ，其*csm* = 0

**确保：**对于*UEm* ，其*csm* = 0

找到最近的节点和集群特征*CFi*

**如果***UEm* 处于*CFi* 的覆盖范围内**，那么**

将*UEm* 加入*CFi*

# 否则

创建*CFj* ，并将*UEm* 加入*CF*中。*j*

**如果**当前节点的CF数量超过了阈值**，那么**

将该节点分成两个节点

将原节点的CF分配给新节点

# end if end if

更新当前节点的路由中的CFs

顺天 簇 在 叶子 作为 *集群* =

*clus*1 *, ..., clusn , ..., clusN*

*{ }*

为每个集群索引*clus*找到最小包围圈，并分配一个无人机。

在[[27](#_bookmark49)]中，作者提出了一种基于价值的强化学习算法，名为Q-learning。代理人可以在状态*s*(*t*)中获得奖励，并通过Q-learning的每一步来选择行动*a*。代理人建立了一个Q-Table

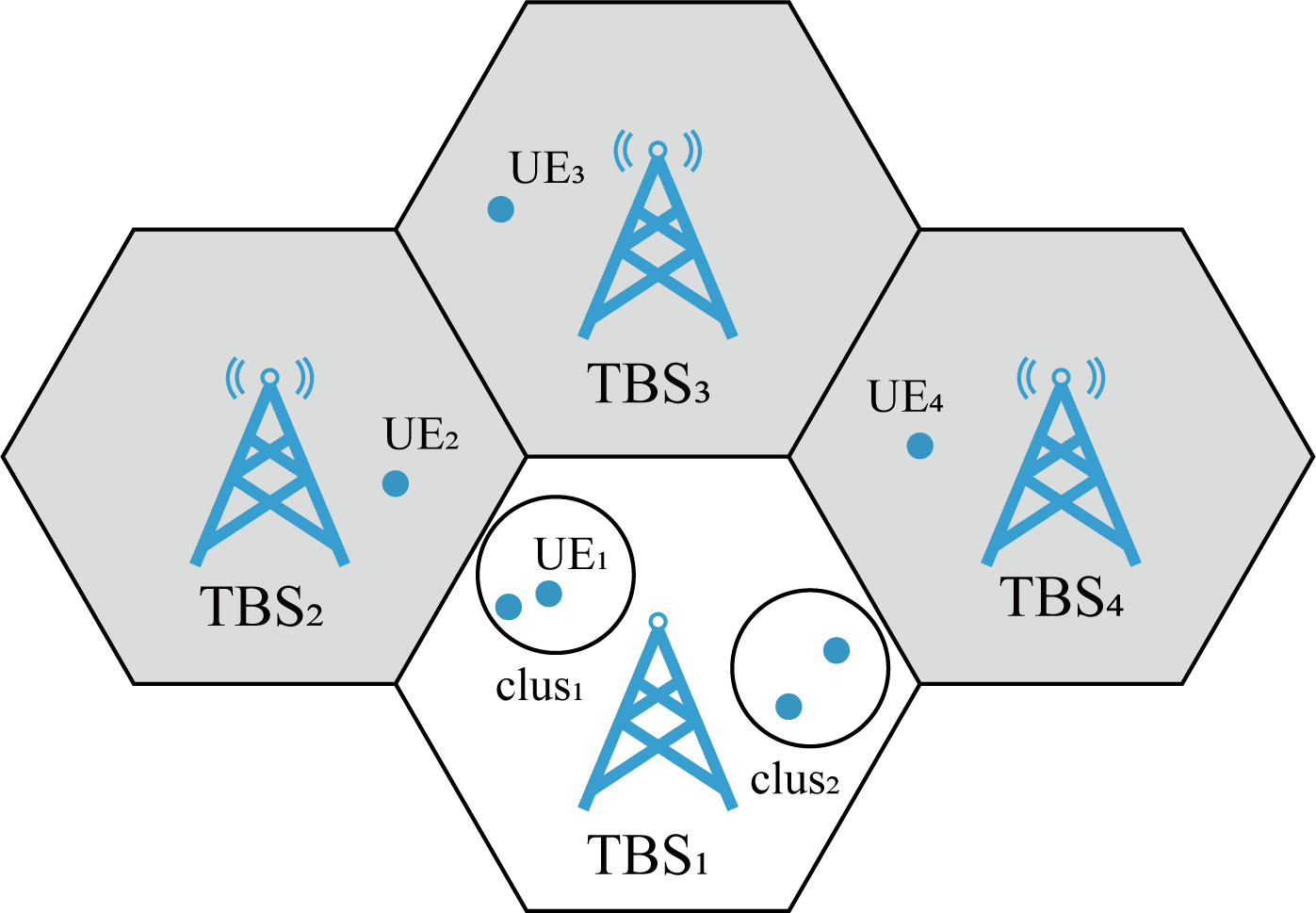


图7.无人机的区域划分

其中*an (t*)是无人机*n* 的动作，定义为*an (t*) 0*,* 1*,* 2*,* 3*,* 4*,* 5*,* 6*，*分别代表向前、向后、向左、向右、向高、向低或盘旋移动。在代理人从*n* *(t*) 0*,* 1*,* 2*,* 3*,* 4*,* 5*,* 6中选择*一个*动作后，代理人的状态变成*S*(*t+1*)，并在不同的状态下重复选择动作的步骤，直到满足表示无人机耗尽电池的步骤限制。

*∈ { }*

*∈ { }*

正如我们在[第三节](#_bookmark1)中所讨论的，*Mn* 和*速率n* 分别是连接用户的数量和用户的速率，由无人机*n* 。当UE的速率超过我们设定的*速率*阈值*th ，Mn* 会增加。有了*速率*阈值*th* ，我们可以扩大空中基站的覆盖范围，提供可靠的无线通信服务。奖励应该被解决，以增加用户的数量和速率。

那么我们可以用以下方式表示奖励函数

(0*, Mn (t*) *≥ Mtotal n 。*

*Mn*(*t*)

来存储状态、行动和Q值，表示为

*q*(*st , a*)，最后选择行动以达到最大值

*rn (t*) =

*Mtotaln - 1, 其他。*

(20)

基于Q-Table的奖励。在此，我们假设无人机部署问题是一个马尔可夫决策过程（MDP），未来状态*st*+1 只由当前状态*st* 决定，那么该问题可以用Q-Learning来处理。

模型中的状态是。

*S*(*t*) = *{s*1 *(t*)*, s*2 *(t*)*, ..., sn (t*)*, ..., sN* (*t)}。* (17)

其中，*sn （t*）表示代理人的状态*n* ，无人机在这里是代理人。代理人的数量*N*将与BIRCH中的集群数量有关。对于每个无人机，状态包括无人机在场景中的直角坐标，其表述为

其中*Mtotal n*是无人机服务的UE总数*n* 。无人机将探索在训练时用可靠的无线通信服务更快地覆盖更多的UE。

在聚类和工作场所划分之后，来自其他代理人的干扰将减少，无人机可以有效地进行部署。这里的代理人是无人机，彼此独立。每次代理人观察环境时，他们都会根据他们的Q-表分别采取行动。在这里，通过Q表选择行动，我们用epsilon贪婪策略设计为

(*1-ϵ+є n , an* = arg maxQ(*sn , an* )*。*

*sn (t*) = *{xn (t*)*, yn (t*)*, zn (t)}。* (18)

*πn (sn |an* ) =

*k*

*єn , 其他。*

(21)

用户与无人机之间的距离可以在每次无人机移动时计算出来，同时可以计算奖励。

当获得无人机*n* "位置 "的信息时，代理人*n* 可以根据*sn (t*)在行动空间中选择一个行动。无人机的行动空间可以表示为

*A*(*t*) = *{a*1 *(t*)*, a*2 *(t*)*, ..., an (t*)*, ..., aN* (*t)}* (19)

*k*

其中*πn (sn an* )是*sn* 中行动的可能性，*k*是行动的数量。由于代理人对环境没有足够的信息，我们希望它最初是随机探索的。

*|*

因此，我们使用一个衰减参数，即ϵ*n* = *de ϵn* ，其中*de*是我们设定的一个常数，ϵ将每集衰减。此外，我们将确定一个常数，即衰变的阈值

*∗*

**算法2** 基于Q-learning的无人机运动

**要求：**环境和Q-表

将代理部署在按预处理划分的区域内

**确保：**为每一集初始化*s*，为每一集的步骤**的**无人机**做**

表一

环境参数

**如果 ϵ***n*

大于 ϵ

*nmin*

# 然后

*an*  arg max *Q*(*sn , an* )

*←*

|  |  |
| --- | --- |
| 参数 | 价值 |
| *ƒc* | 2Mhz |
| *铂金* | 4kW |
| *tt* | 1 |
| *叁* | 1 |
| *Bn* | 0.18Mhz |
| *N*0 | -174dBm/H7 |
| *ht* | 100m |
| *β* | 1 |
| *α* | 1 |
| *µlos* | 3 |
| *µnlos* | 23 |
| *n* | 5 |

# 否则

在行动空间中随机生成*一个n*

# 结束 如果

观察*r*和*s′*

**如果***S′ n*等于其他UAVs的*S′* **，那么**

*s′ n sn*

*←*

# 结束 如果

观察*r*和*s′*

*Q*(*s, a*) ← *Q*(*s, a*) + *α{r* + ц max[*Q*(*s′ , a′* ) *- Q*(*s, a*)]*}.*

*s ← s′*

*ϵn ϵn de*

*←*

# 结束

参数，ϵ*nmin* 。一旦ϵ*n* 下降到ϵ*nmin* 以下，衰减参数将被设定为ϵ*nmin* ，因为

表二

Q-learning的参数

|  |  |
| --- | --- |
| 参数 | 价值 |
| *g* | 0.9 |
| *є* | 0.9 |
| *德* | 0.999 |
| *єmin* | 0.01 |
| *Lr* | 0.5 |

而无人机的初始坐标是在指定区域内随机的。我们将无人机的工作地点设定为城市，因此

*ϵn* =

(ϵn

*, ϵn*

*≥ ϵnmin,*

(22)

恒定值*α*=1和β=1。无人机的LOS链路和NLOS链路的衰减系数为*µlos* = 3，*µnlos* = 23。

*ϵnmin, ϵn < ϵnmin。*

然后，代理将在一个不熟悉的环境中随机探索，然后充当epsilon贪婪。所有用户将根据上述协议连接到传输设备。奖励将根据用户的数量和速率更新。此外，我们在这里增加了碰撞检测。如果一个无人机的下一个状态与*s′ n*相同，它将在下一刻被确定为一个碰撞。那么后一个的行动将改变为停留在这个时刻，而前一个继续行动。碰撞检测后，代理将根据最终的奖励*r*和状态*s′* ，更新Q表和状态。如上所述，由于代理独立工作，奖励是各自的。我们将整个系统的性能表述为

Σ

*rsys* = *rn 。* (23)

*n*

通过系统的奖励*rsys* ，我们可以评估系统的性能，并使所有UE都得到服务。

1. 模拟结果

我们将仿真环境设定为5200\*5200\*5200的网格世界。在这个模拟中，有4个状态为0*、*1*、*1*、*0的TBS。TBSs和UAVs的覆盖范围分别为2400、1200。[表一](#_bookmark9)显示了环境和基站的参数。我们假设TBSs和UAVs的中心频率为*fc* = 2Ghz，发射器功率为*Pt* = 4kW。TBSs的发射机增益和UEs的接收增益为*Gt* = *Gr* = 1。所有通道的带宽为*Bn* = 0*.*18Mhz。同时，我们假设所有TBS的高度为*ht* = 100m。

*{ }*

分别。此外，用于计算噪声的功率谱密度为174dBm/HZ。

同时，我们将Q-learning的参数设置[为表二](#_bookmark10)。学习率为*lr* = 0*.*5，奖励衰减为ц = 0*.*9。此外，ε-greedy算法的初始参数为ϵ = 0*.*9，衰减参数为*de* = 0*.*999。ϵ的阈值为 ϵ*min* = 0*.*01。然后设置*ep*=10000，模拟10000个事件。

*-*

图[8](#_bookmark11)描绘了当TBS的状态为0,1,1,0时UE的集群。在模拟中，有400个UE。由于基站在灾难发生前覆盖了所有的人，所以用户在基站的覆盖区域内围绕基站分布。我们假设用户在地理上以泊松点过程（PPP）分布在TBS周围。然后我们结合UEs在每个基站的分布情况，如图[8](#_bookmark11)所示。图中的黑点代表灾后由TBS服务的UE，四个褐红色的点是TBS。因为如上所述，TBS的状态是0*,* 1*,* 1*,* 0，所以需要无人机提供紧急无线服务的用户主要分布在TBS1 和TBS4 的覆盖范围内。也就是说，无人机的工作场所是两个260\*260\*260的网格空间，而TBS则在其余部分为用户服务。由于环境的比例是1:20m，无人机将在两个5200m\*5200m的区域内工作。如上所述，无人机的传输距离为1500米。为了确保无人机覆盖整个指定区域，我们用BIRCH对无人机覆盖的UE进行聚类，未连接的UE被分为八个聚类。图中由彩色点填充的圆圈是无人机为未连接的UE服务的区域。而这些彩色的点是BIRCH之后的集群。集群的数量是8个，因此，我们将部署8架无人机来支持TBSs，以便

*{ }*

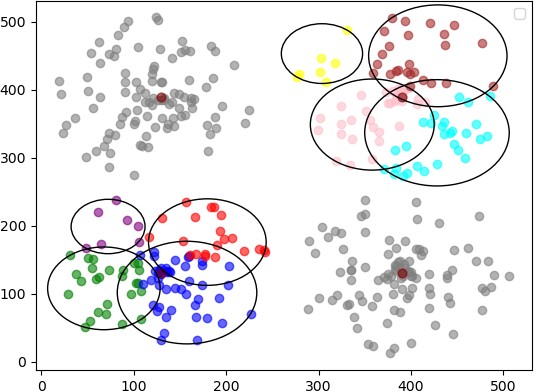
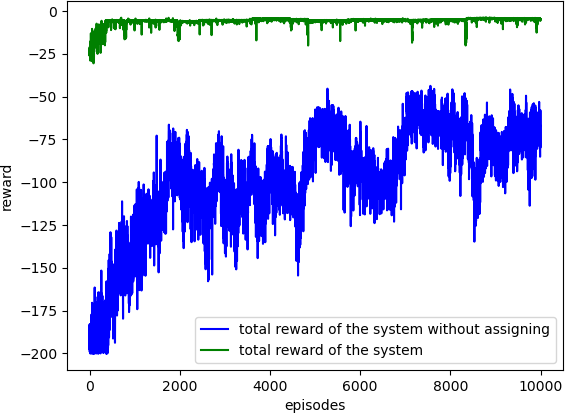
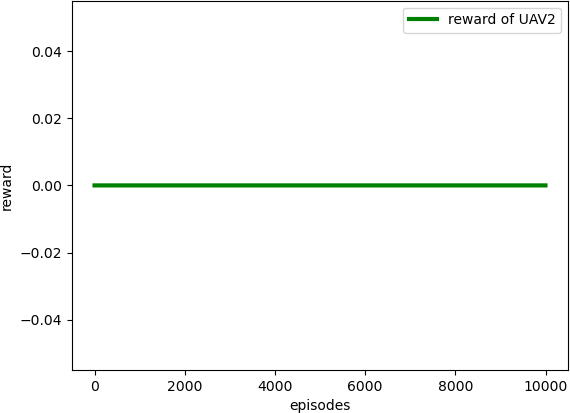
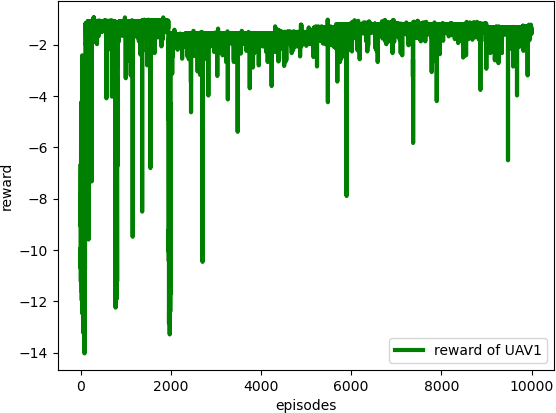
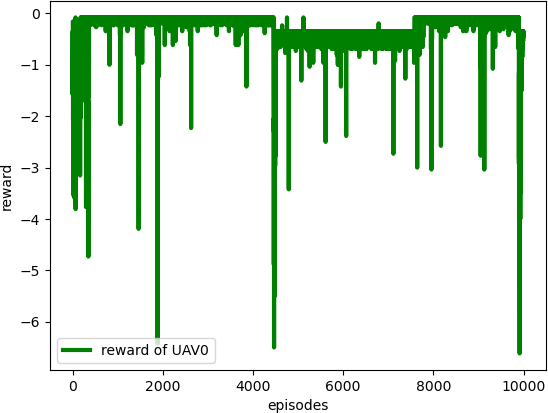
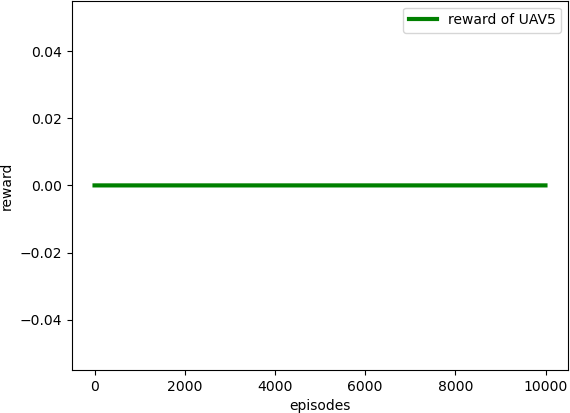
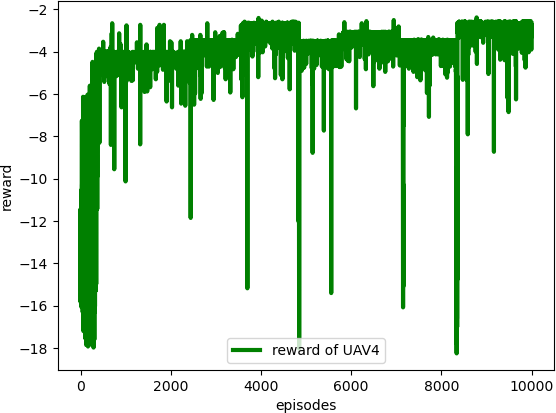
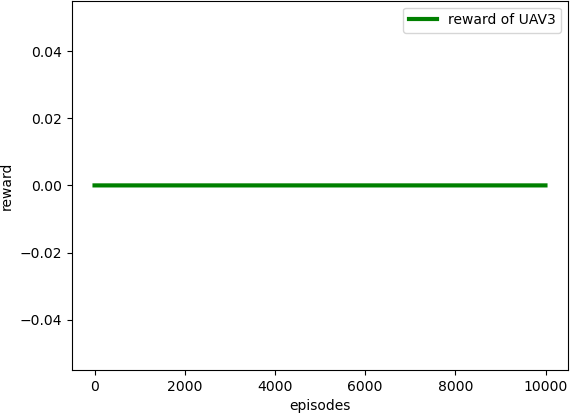
 

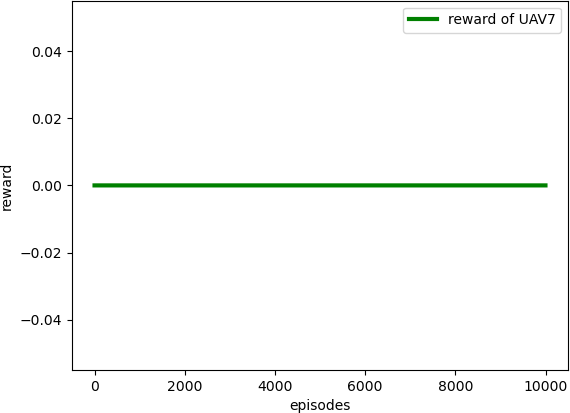
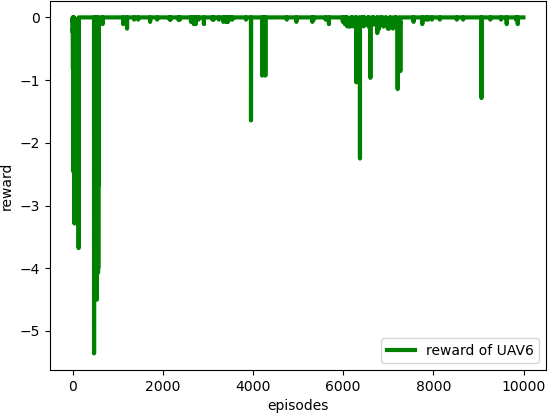
图8.使用Birch的用户集群



* 1. 代理人0 (b) 代理人1 (c) 代理人2



(d) 代理人3 (e) 代理人4 (f) 代理人5



(g) agent6  (h) agent7 图9.每个代理人的奖励

覆盖这个灾后地区的所有UE。

图[9](#_bookmark18)显示了每个代理人的奖励。图[9(a](#_bookmark12)), 图9([b](#_bookmark13)), 图[9(c](#_bookmark14)), 图9([d](#_bookmark15)), 图9([e)](#_bookmark16), 图9([f](#_bookmark17)), 图[9(g)](#_bookmark19)和图[9(h)](#_bookmark20)是*无人机*0 , *无人机*1 , *无人机*2 , *无人机*3 , *无人机*4 ，*无人机*5 的奖励。

分别是*无人机*6 ，和*无人机*7 。在对没有服务的UE进行聚类时，距离的阈值等于无人机的最大覆盖范围。而无人机的初始位置在其指定区域内是随机的。因此，在实施BIRCH分配无人机时，有些无人机会碰巧覆盖其工作地点的所有UE，如*无人机*2 、*无人机*3 、*无人机*5 、*无人机*7 。而其他无人机，如*UAV*0 , *UAV*1 , *UAV*4 *,UAV*6 , 会很容易地覆盖所有的UE，绿色曲线会在训练步骤后收敛，因为它们被分配到符合其最大覆盖范围的合适工作场所。

我们在[第四节](#_bookmark7)中提到了系统的奖励*rsys* ，我们将其绘制在图[10](#_bookmark21)中。它显示了系统每集的奖励。随着时间的推移，绿色曲线和蓝色曲线的奖励都在增加。但与没有分配BIRCH的蓝色曲线相比，分配了BIRCH的绿色曲线会稳定地获得更高的奖励。经过几步的学习，无人机将完成覆盖所有未服务的UE的任务，其可靠率高于我们之前设定的阈值。

图10.系统的总回报

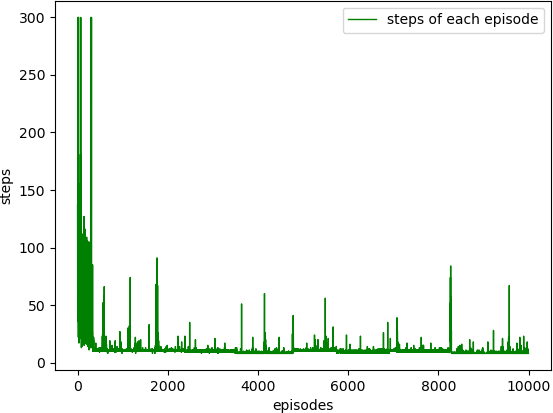


图11.每集的步数

迅速地。

此外，图[11](#_bookmark22)显示了系统每集覆盖所有UE的平均步骤数。因为我们设定的无人机的电池是300，低于300的步数意味着在这些情节中覆盖了所有的UE。经过一段时间的训练，系统可以在电池耗尽前覆盖所有用户。此外，经过BIRCH的聚类和区域划分，我们减少了每个无人机的工作量。无人机覆盖所有未连接的UE的步骤迅速减少。无人机将为所有的UE服务，并在一些步骤的学习后迅速完成任务。

1. 结论

灾难发生后，稳定的无线通信服务对救援是至关重要的；然而，一些TBS总是坠毁，而地面应急设施不能立即部署。随着无人机的发展，由于无人机的敏捷性，我们可以将无人机部署为空中基站。为了最大限度地覆盖灾区的灾民，我们提出了一种基于Q-learning的无人机移动方案，与此同时，我们还提出了一种基于Q的学习方案。

现有的可用TBS为所有受害者提供无线通信。在Q-learning之前，我们使用基于空中基站覆盖范围的聚类算法对这些没有服务的UE进行分类，缩小无人机的工作范围，然后根据聚类的结果动态部署无人机。然后，无人机可以在Q-learning中更快、更容易地获得最大的回报。通过模拟，无人机可以在灾后及时覆盖所有没有被TBS服务的用户。

鸣谢

这项工作得到了JSPS KAKENHI拨款号JP19K20250、JP20F20080和JP20H04174的部分支持。

日本文部科学省优秀青年研究者倡议（LEADER），以及日本科技部PRESTO拨款号JP- MJPR21P3，日本。董绵雄是通讯作者。

参考文献

1. J.Xu, K. Ota, and M. Dong, "Big data on the fly:Uav-mounted mobile edge computing for disaster man- agement," *IEEE Transactions on Network Science and* *Engineering*, 2020.
2. J.Xu, K. Ota, and M. Dong, "Fast deployment of emer- gency fog service for disaster response," *IEEE Network*, vol. 34, no. 6, pp. 100-105, 2020.
3. A.Merwaday，"基于随机几何学的密集het- erogeneous网络的容量、移动性和能源效率分析"，2016。
4. J.Xu, K. Ota, and M. Dong, "Fast networking for disaster recovery," *IEEE Transactions on Emerging Topics in* *Computing*, 2017.
5. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Gior-dano, A. Garcia-Rodriguez, and J. Yuan, "Survey on Uav cellular communications:Practical aspects, standardiza- tion advancements, regulation, and security challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no.4, pp. 3417-3442, 2019.
6. H.Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou,

E.Almaita, I. Khalil, N. S. Othman, A. Khreishah, and

M.Guizani, "Unmanned aerial vehicles (uavs):A survey on civil applications and key research challenges," *Ieee* *Access*, vol. 7, pp. 48 572-48 634, 2019.

1. M.Mozaffari, W. Saad, M. Bennis, Y.-H.Nam, and

M.Debbah, "A tutorial on uavs for wireless networks:应用、挑战和开放的问题，"*IEEE通信调查与教程*，第21卷，第。3, pp. 2334-2360, 2019.

1. B.Li, Z. Fei, and Y. Zhang, "Uav communications for 5g and beyond:Recent advances and future trends," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2241-2263, 2018.
2. J.Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of uav-mounted mobile base stations, " *IEEE Communications Letters*, vol. 21, no.3, pp. 604-607, 2016.
3. 亚马逊。(2016, 12) Amazon prime air的第一个客户交付。[在线].Available: [https://www.youtube.com/ watch?v=vNySOrI2Ny8](https://www.youtube.com/watch?v=vNySOrI2Ny8)
4. J.斯图尔特。(2014, 08) 谷歌在项目机翼试验中测试无人机送货。[在线]。Available: [https://www.bbc.](https://www.bbc.com/news/technology-28964260) [com/news/technology-28964260](https://www.bbc.com/news/technology-28964260)
5. I.K. Durrani.(2016, 8) 诺基亚和EE--用无人机推动技术的极限。[在线]。可用[：https://www.nokia.com/blog/nokia-ee-pushing-limi](https://www.nokia.com/blog/nokia-ee-pushing-limits-technology-drones/) [ts-technology-drones/](https://www.nokia.com/blog/nokia-ee-pushing-limits-technology-drones/)
6. S.李维。(2013, 06) 谷歌将如何利用高空气球向内陆地区提供互联网。[在线]。见[：https://www.wired.com/2013/06/google](https://www.wired.com/2013/06/google_internet_balloons/) [internet balloons/](https://www.wired.com/2013/06/google_internet_balloons/)
7. 高通公司和AT&T将在蜂窝网络上试用无人机，以加速大范围的部署。[在线]。Available: [https://www.qualcomm.com/news/releases/2 016/09/06/qualcomm-and-att-trial-drones-cellular-netwo](https://www.qualcomm.com/news/releases/2016/09/06/qualcomm-and-att-trial-drones-cellular-network-accelerate-wide-scale) [rk-accelerate-wide-scale](https://www.qualcomm.com/news/releases/2016/09/06/qualcomm-and-att-trial-drones-cellular-network-accelerate-wide-scale)
8. Q.Feng, E. K. Tameh, A. R. Nix, and J. McGeehan, "Wlcp2-06:城市环境中空对地无线电传播的视线可能性建模，" in *IEEE Globecom 2006*.IEEE，第1-5页。
9. A.Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal lap altitude for maximum coverage," *IEEE Wireless Com-* *munications Letters*, vol. 3, no. 6, pp.569-572, 2014.
10. M.Mozaffari, W. Saad, M. Bennis, and M. Debbah, "云中的无人机小电池。设计、部署和性能分析"，在*2015年IEEE全球通信会议（GLOBECOM）*上。IEEE，2015年， ，第1-6页。
11. I.Bor-Yaliniz和H. Yanikomeroglu，"牧场异质性的新前沿。Multi-tier dron-cells," *IEEE Com-* *munications Magazine*, vol. 54, no. 11, pp. 48-55, 2016.
12. E.Kalantari, H. Yanikomeroglu, and A. Yongacoglu, "On the number and 3d placement of drone base stations in wireless cellular networks," in *2016 IEEE 84th Vehicular Technology Conference（VTC-Fall）*.IEEE，2016，第1- 6。
13. M.Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Ef- ficient deployment of multiple unmanned aerial vehicles for optimal wireless coverage," *IEEE Communications* *Letters*, vol. 20, no. 8, pp. 1647-1650, 2016.
14. E.Kalantari, M. Z. Shakir, H. Yanikomeroglu, and

A.Yongacoglu，"Backhaul-aware robust 3d drone place- ment in 5g+ wireless networks," in *2017 IEEE inter- national conference on communications workshops（ICC* *workshops）*。IEEE，2017，pp.109-114。

1. Y.Zeng and R. Zhang, "Energy-efficient uav communi- cation with trajectory optimization," *IEEE Transactions on Wireless Communications*, vol. 16, no. 6, pp. 3747- 3760, 2017.
2. M. Alzenad, A. El-Keyi, F. Lagum, and

H.Yanikomeroglu, "3-d placement of an unmanned aerial vehicle base station (uav-bs) for energy-efficient maximal coverage," *IEEE Wireless Communications* *Letters*, vol. 6, no.4, pp. 434-437, 2017.

1. X.Li, "Deployment of drone base stations for cellular communication without apriori user distribution informa- tion," in *2018 37th Chinese Control Conference（CCC）*.IEEE, 2018, pp.7274-7281.
2. X.Liu, Y. Liu, and Y. Chen, "Reinforcement learning in multiple-uav networks:Deployment and movement design," *IEEE Transactions on Vehicular Technology*, vol.68, no.8, pp.8036-8049, 2019.
3. S.A. Hoseini, J. Hassan, A. Bokani, and S. S. Kan- here, "Trajectory optimization of flying energy sources using q-learning to recharge hotspot uavs," *arXiv preprint* *arXiv:2003.12258*, 2020.
4. C.J. Watkins和P. Dayan, "Q-learning," *Machine learning*, vol. 8, no.3-4, pp. 279-292, 1992.

**赵世业**于2019年在中国北京邮电大学获得工学学士学位，同年在日本室兰工业大学的ENeS实验室做研究。目前他正在日本室兰工业大学攻读硕士学位。

**太田薰**出生在日本会津若松市。2008年，她在美国俄克拉荷马州立大学获得计算机科学硕士学位。

2006年和2012年分别获得日本会津大学计算机科学与工程专业的学士和博士学位。薰目前是室兰工业大学科学和信息学系的副教授和教育、文化、体育、科学和技术部（MEXT）的优秀青年研究员。

日本。2010年3月至2011年3月，她在加拿大滑铁卢大学做访问学者。此外，2012年4月至2013年4月，她还在日本东北大学担任日本科学促进会（JSPS）的研究员。Kaoru是2017年IEEE TCSC早期职业奖、2018年第13届IEEE ComSoc亚太青年研究员奖、2020年N2Women的获得者。计算机网络和通信领域的新星，2020年KDDI基金会鼓励奖，以及2021年IEEE札幌青年专业人士最佳研究员奖。她是Clarivate Analytics 2019年、2021年高被引研究员（Web of Science），2021年被选为JST-PRESTO研究员，2022年成为EAJ研究员。

**董绵雄**在日本会津大学获得计算机科学与工程的学士、硕士和博士学位。他是日本室兰技术学院的副院长和教授。他曾是日本会津大学计算机科学与工程学院的JSPS研究员，并于2010年4月至2011年8月在加拿大滑铁卢大学BBCR小组做访问学者，由JSPS优秀青年海外访问计划支持。董博士被选为外国人

2011年被NEC C&C基金会评为研究员（全日本共有3人获得）。他是2017年第12届IEEE ComSoc亚太青年研究员奖、2018年Funai研究奖、2018年NISTEP研究员（日本仅有的11人之一），以表彰其在科学和技术方面的重大贡献，2021年日本文部省的青年科学家奖，2021年IEIEC的SUEMATSU-Yasuharu奖，2021年IEEE TCSC中等职业奖。他是Clarivate Analytics 2019年、2021年的高被引研究员（Web of Science）和EAJ的外籍研究员。