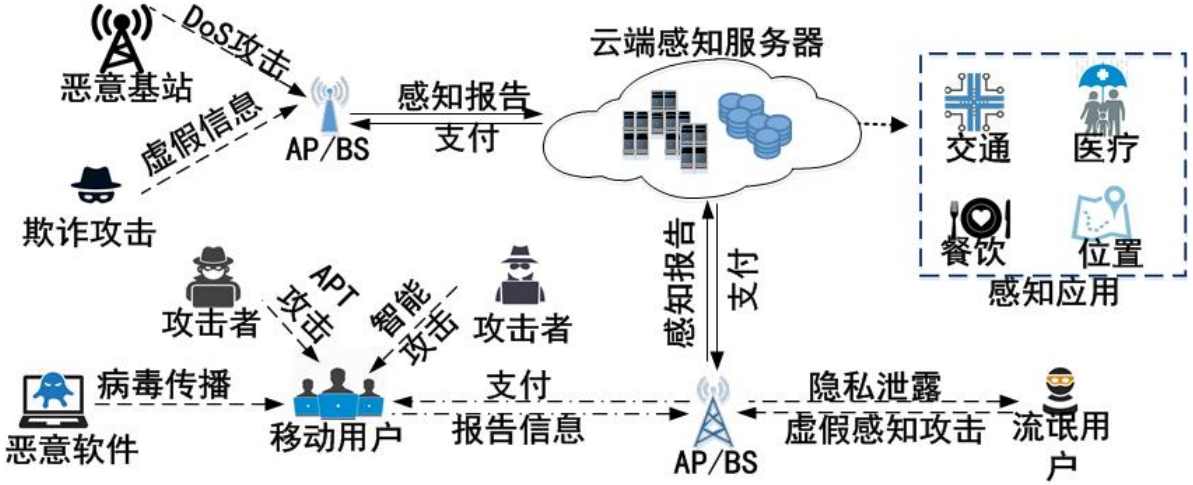
基于深入强化学习的移动群智感知机制

肖亮
厦门大学

第二届网络空间安全喻园青年科学家论坛
武汉
2018年5月5日

网络安全博弈与机器学习研究



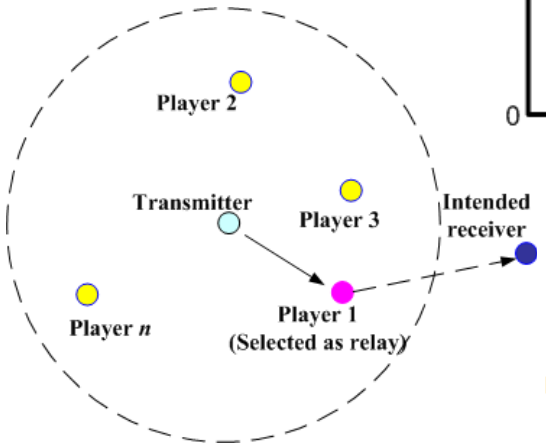
网络攻防博弈	物理层认证	抗敌意干扰
APT攻防 病毒检测 认证，抗敌意干扰	物联网，云存储 水下传感网 智能电网，车联网	物联网 水下传感网 车联网
纳什均衡 斯塔克伯格均衡 强化学习	机器学习	协作型非协调抗跳频 强化学习 深度学习
抑制多种网络攻击 提高自主防御能力	提升认证精度 加快认证速度	降低传输时延 节能 提高信号质量

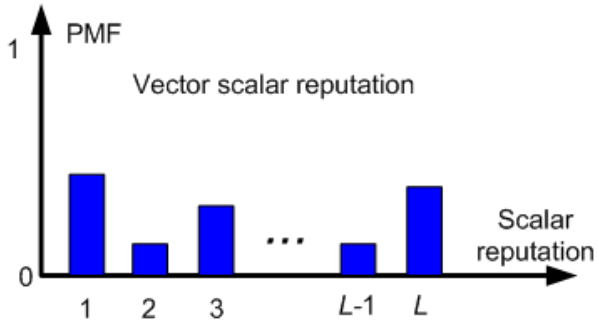
基于博弈机制和强化学习的网络安全方法

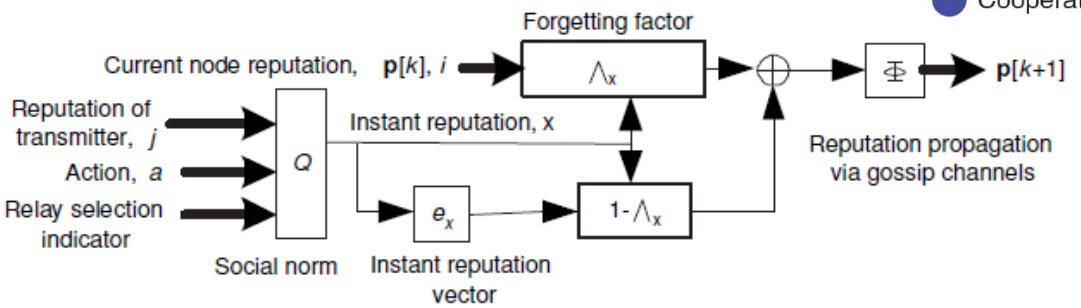
1368 IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 7, NO. 4, AUGUST 2012

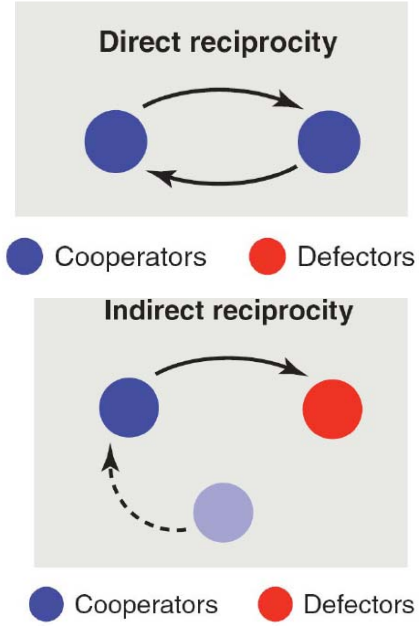
Indirect Reciprocity Security Game for Large-Scale Wireless Networks

Liang Xiao, *Member, IEEE*, Yan Chen, *Student Member, IEEE*, W. Sabrina Lin, *Student Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*





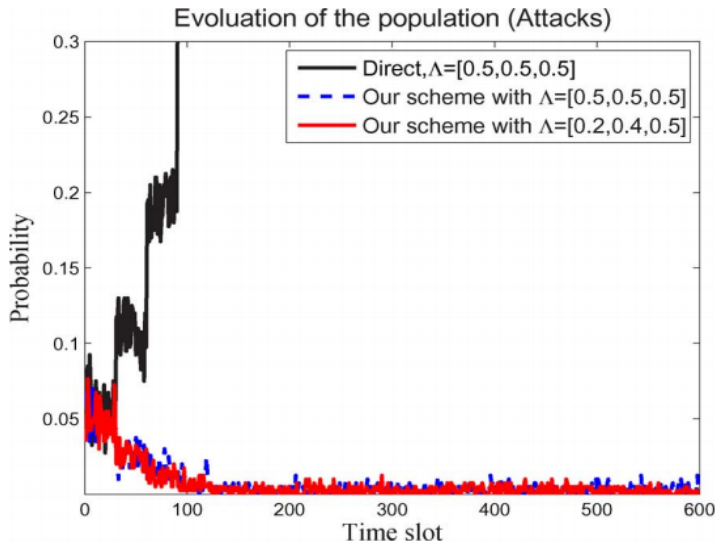




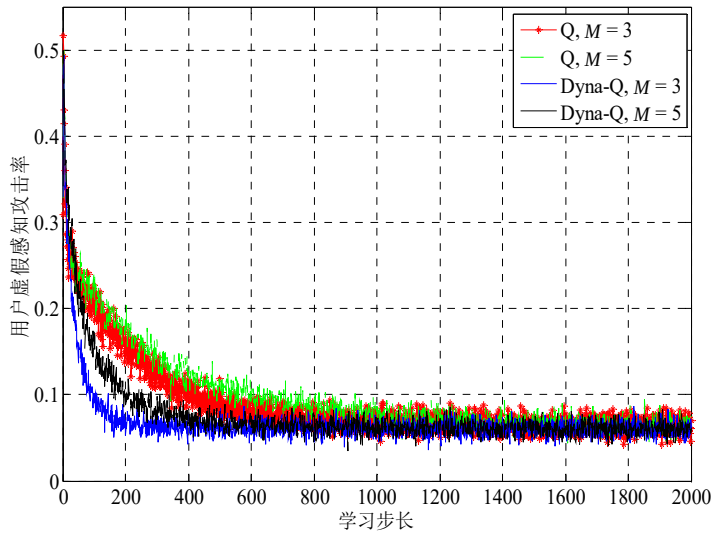
IEEE Trans. Information Forensics & Security

- 采用博弈机制和机器学习设计无线网络安全方案
- 挖掘移动用户的间接关联性，构建网络安全信誉机制，实现主动防御
- 建立无线网络攻防博弈模型，揭示其均衡机理及演化特性

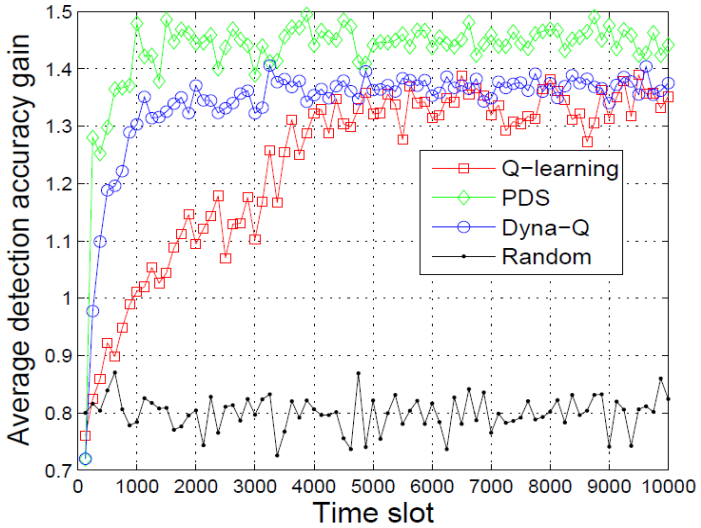
■ 抑制攻击动机，网络被攻击概率降低25%，恶意软件检测精度提高12%



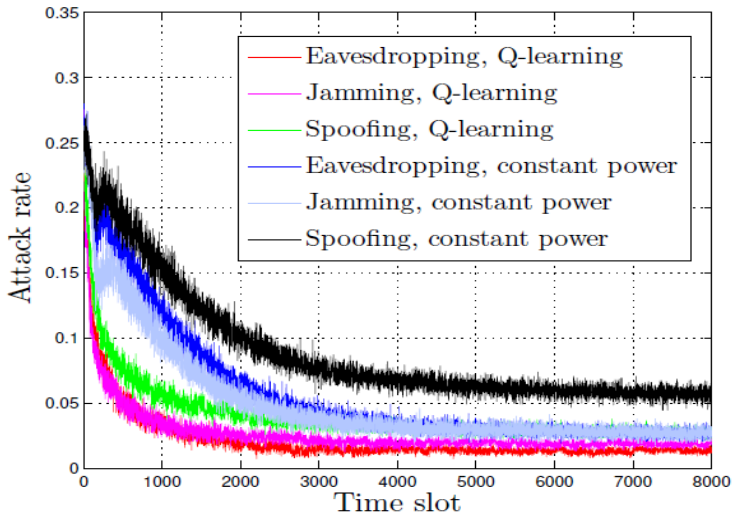
基于信誉机制的无线安全方案



基于云计算的安全移动群智感知



智能化的移动终端病毒检测技术



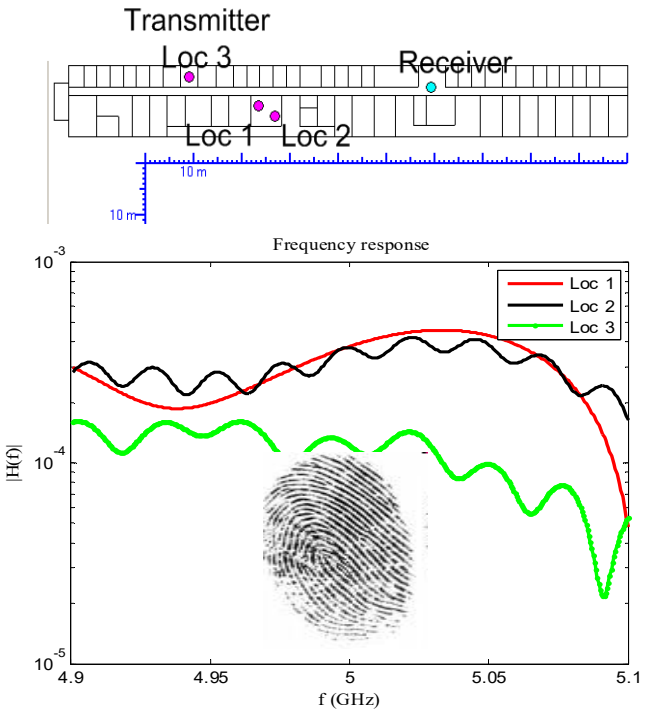
基于多天线系统的抗智能攻击技术

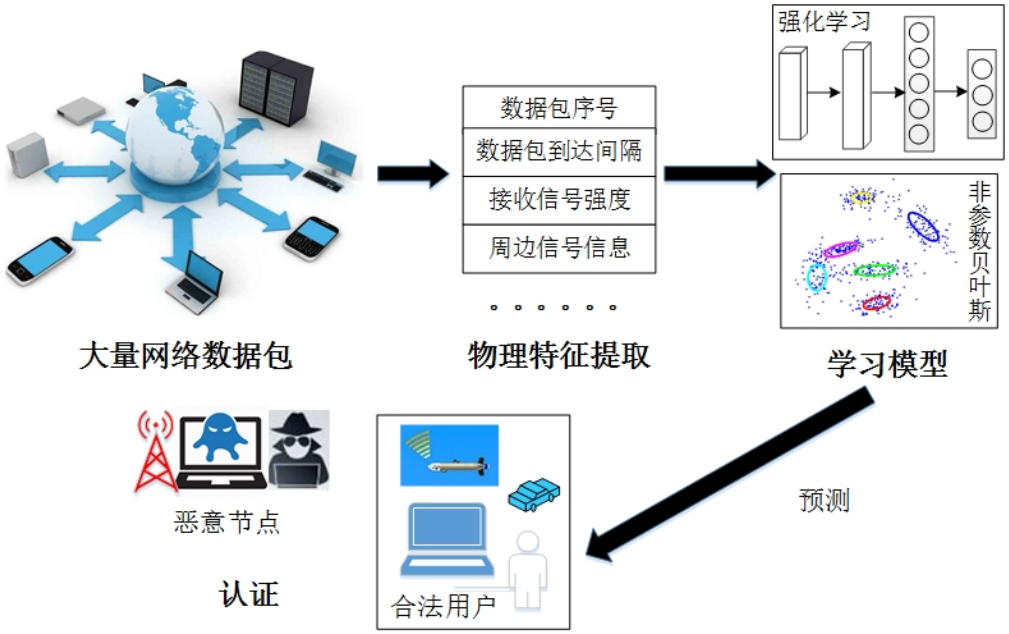
基于博弈的无线网络物理层认证技术

IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 8, NO. 12, DECEMBER 2013

Proximity-Based Security Techniques for Mobile Users in Wireless Networks

Liang Xiao *Senior Member, IEEE*, Qiben Yan, *Student Member, IEEE*, Wenjing Lou, *Senior Member, IEEE*
Guiquan Chen, *Student Member, IEEE*, and Y. Thomas Hou, *Senior Member, IEEE*

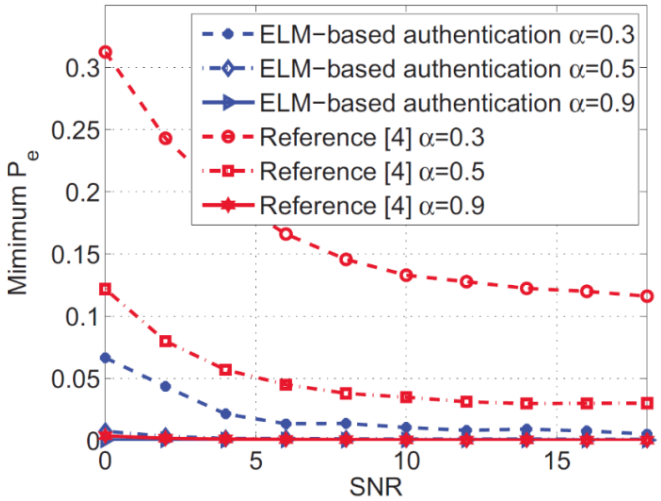


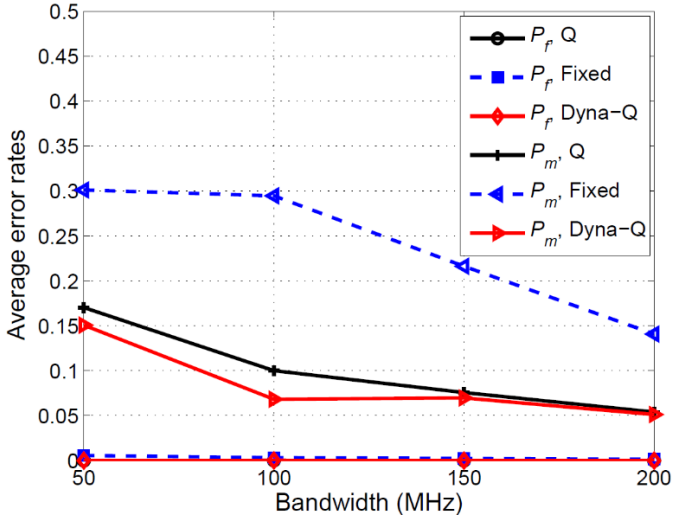


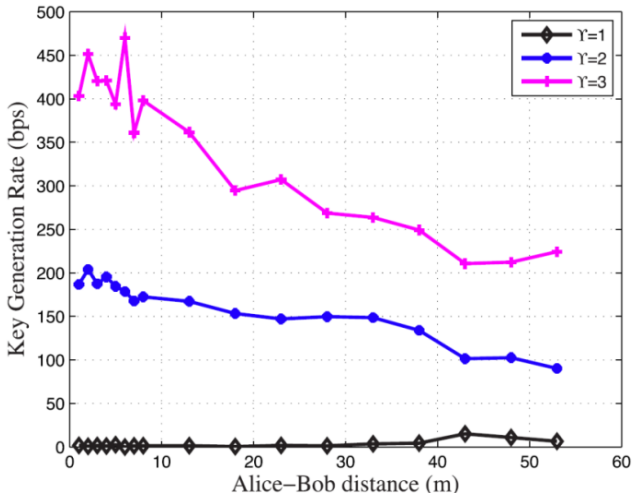
IEEE Trans. Information Forensics & Security

- 提出无线频域信道指纹概念，建立无线认证攻防博弈
- 提出基于强化学习和非参数贝叶斯的无线网络认证方案，保护用户位置隐私

- 保护用户位置隐私，不需预知网络攻击模型
- 电子欺骗检测成功率提高5.4%，虚警率降低63.6%





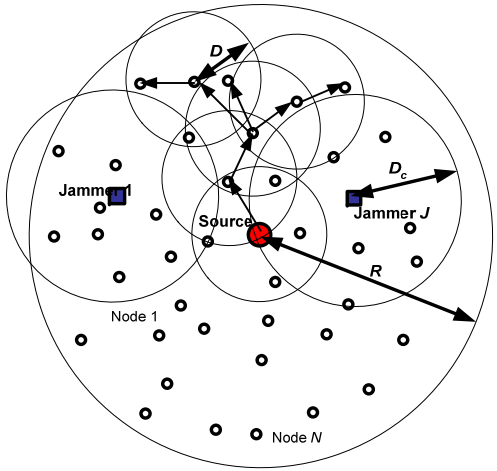


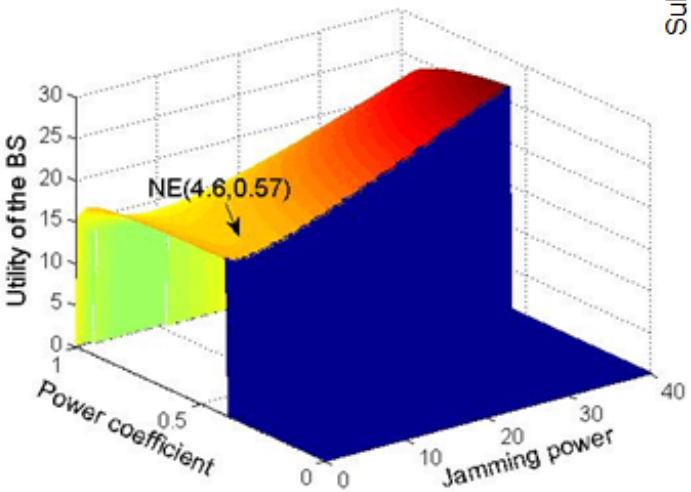
基于博弈论的无线网络敌意干扰攻击研究

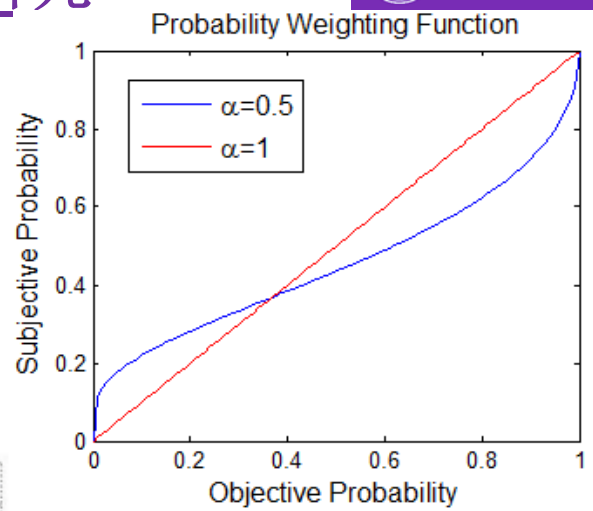
IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 7, NO. 1, FEBRUARY 2012

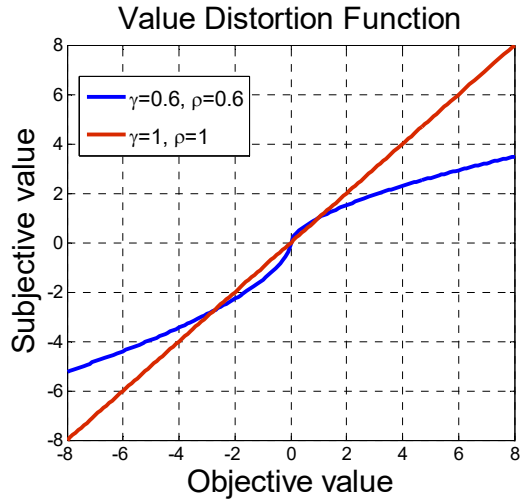
Jamming-Resistant Collaborative Broadcast Using Uncoordinated Frequency Hopping

Liang Xiao, *Associate Member, IEEE*, Huaiyu Dai, *Senior Member, IEEE*, and Peng Ning, *Member, IEEE*





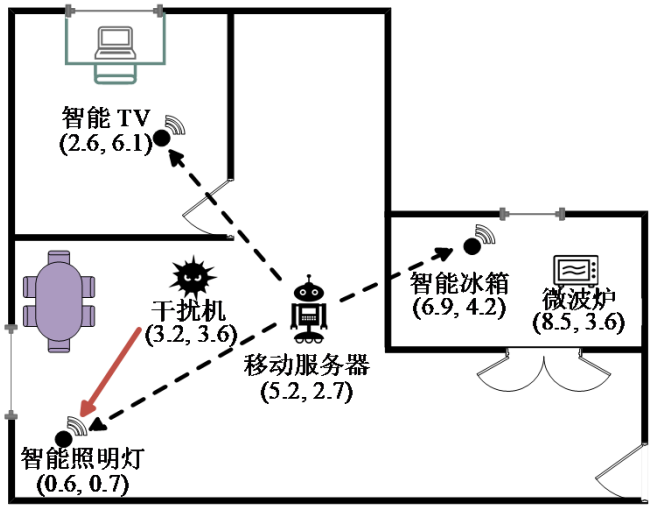


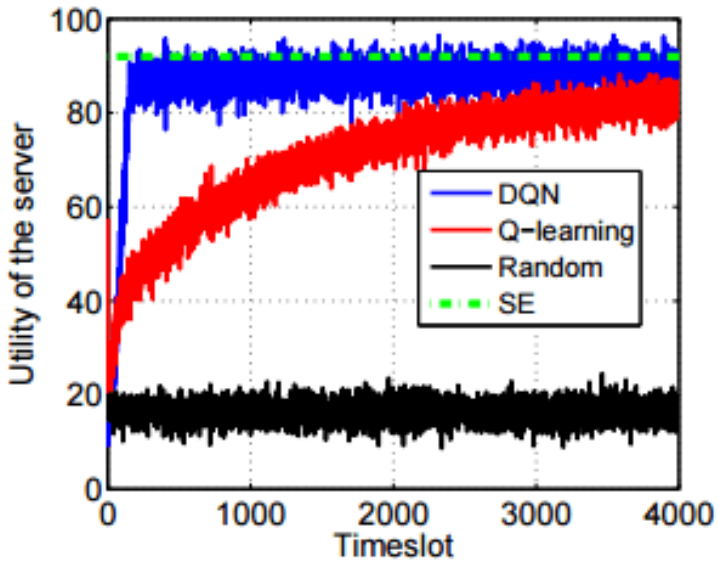


IEEE Trans. Information Forensics & Security

- 针对未知网络攻击模型的场景，提出基于强化学习的无线网络抗干扰传输机制
- 建立无线网络的主观性攻防博弈模型，揭示攻击者的主观性对博弈均衡的影响机理

■ 提高抵御敌意干扰攻击的能力，传输时延降低32%，移动端能耗节约60%，增强无线传输质量



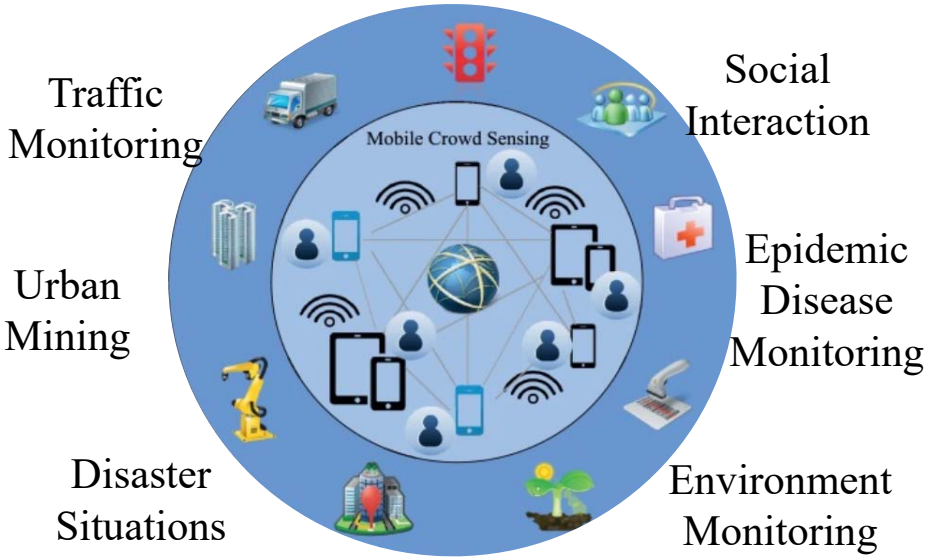


主要内容

- 背景介绍
- 移动群智感知系统的虚假感知攻击
- 基于深度强化学习的移动群智感知系统
- 性能分析
- 总结

移动群智感知(MCS)系统

- 通过移动互联网招募移动用户，分发感知任务，收集感知数据
  - 执行感知任务需要消耗时间、能量、带宽等资源，有隐私泄露的危险
- 有偿式群智感知：服务器通过虚拟货币等方式激励用户参与感知任务



- ✓ 摄像头
- ✓ 音频
- ✓ GPS
- ✓ 压力
- ✓ 环境光线
- ✓ 距离
- ✓ 加速计
- ✓ 指南针
- ✓ 陀螺仪

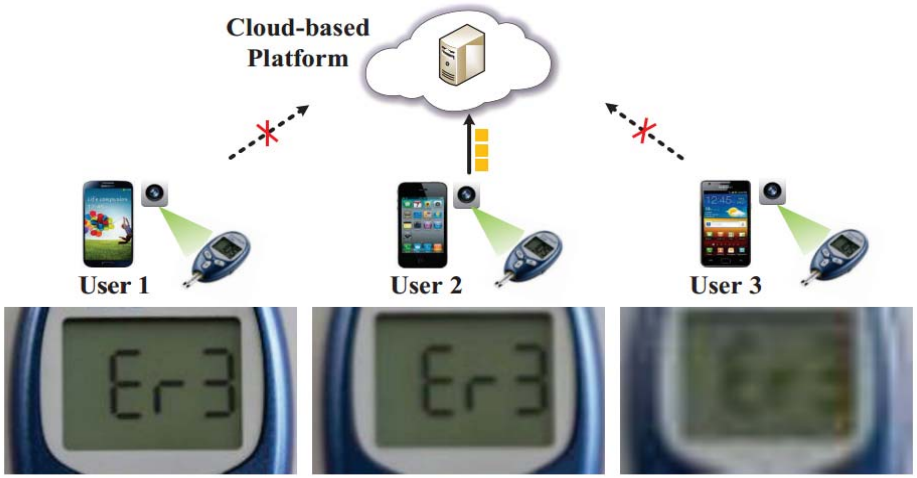
虚假感知攻击

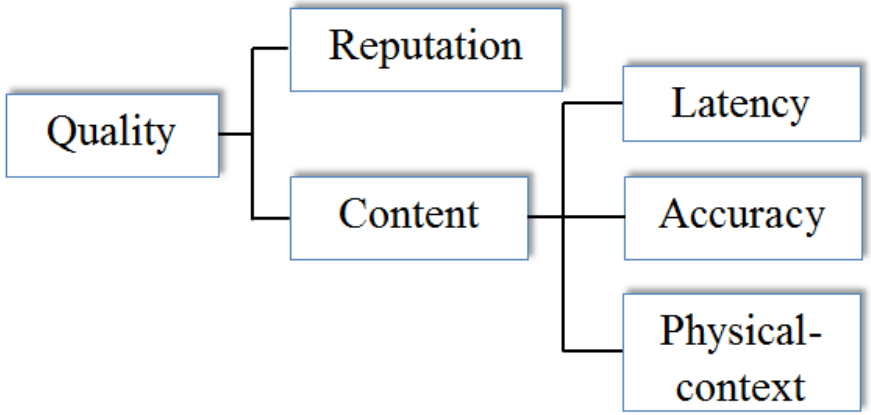
- 自私用户发动虚假感知攻击，发送虚假感知报告
  - 通过智能终端获取自主编程及设备控制能力
  - 降低感知损耗，获取非法收益，保护个人隐私
  - 导致MCS的感知任务失败

https://play.google.com/store/apps/details?id=com.lkr.fakelocation

移动群智分级支付机制

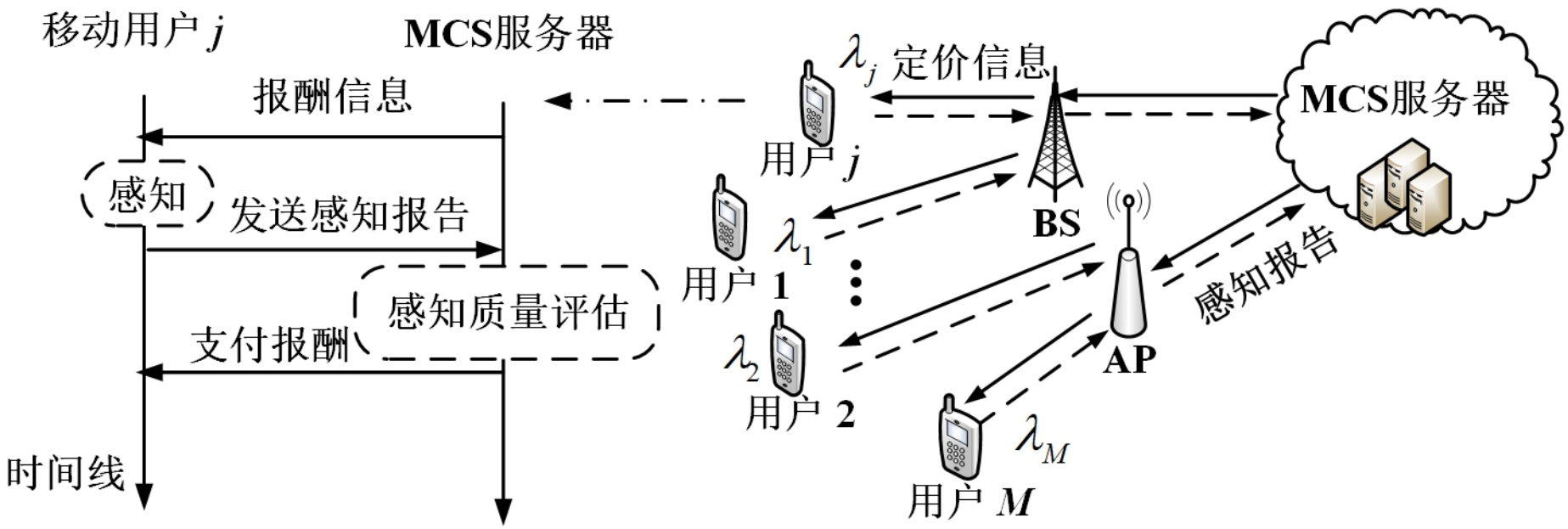
- MCS服务器评估感知质量
  - 基于分类算法[Peng’15]和投票机制[Huang’14]
  - 利用图像失真度等特定指标进行质量评分[Saad’16]
- 采用分级付费等激励机制从动机上抑制自私用户的虚假感知行为





Jin H, Su L, Chen D, et al., " Quality of information aware incentive mechanisms for mobile crowd sensing systems," ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2015.

系统模型

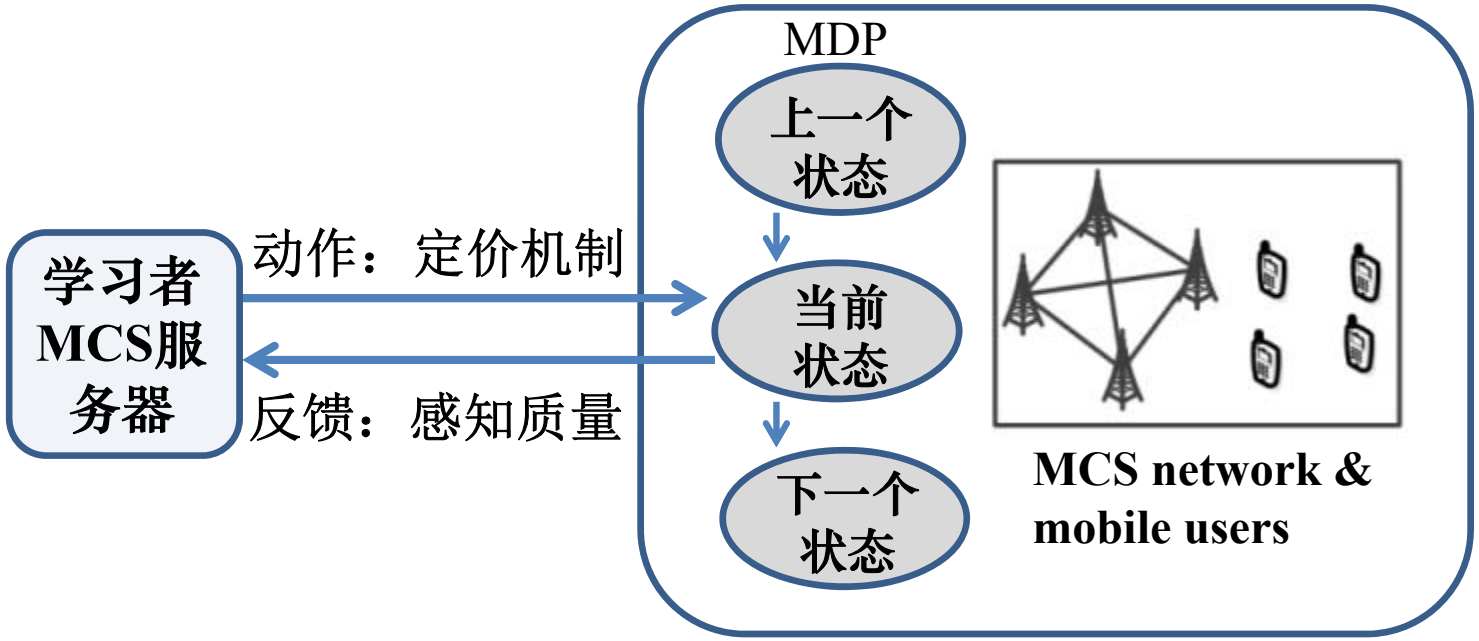

移动用户 j
MCS服务器
报酬信息
感知
发送感知报告
感知质量评估
支付报酬
时间线
λj 定价信息
用户 j
BS
MCS服务器
λ1
用户 1
λ2
用户 2
AP
感知报告
λM
用户 M

安全移动群智感知动态博弈

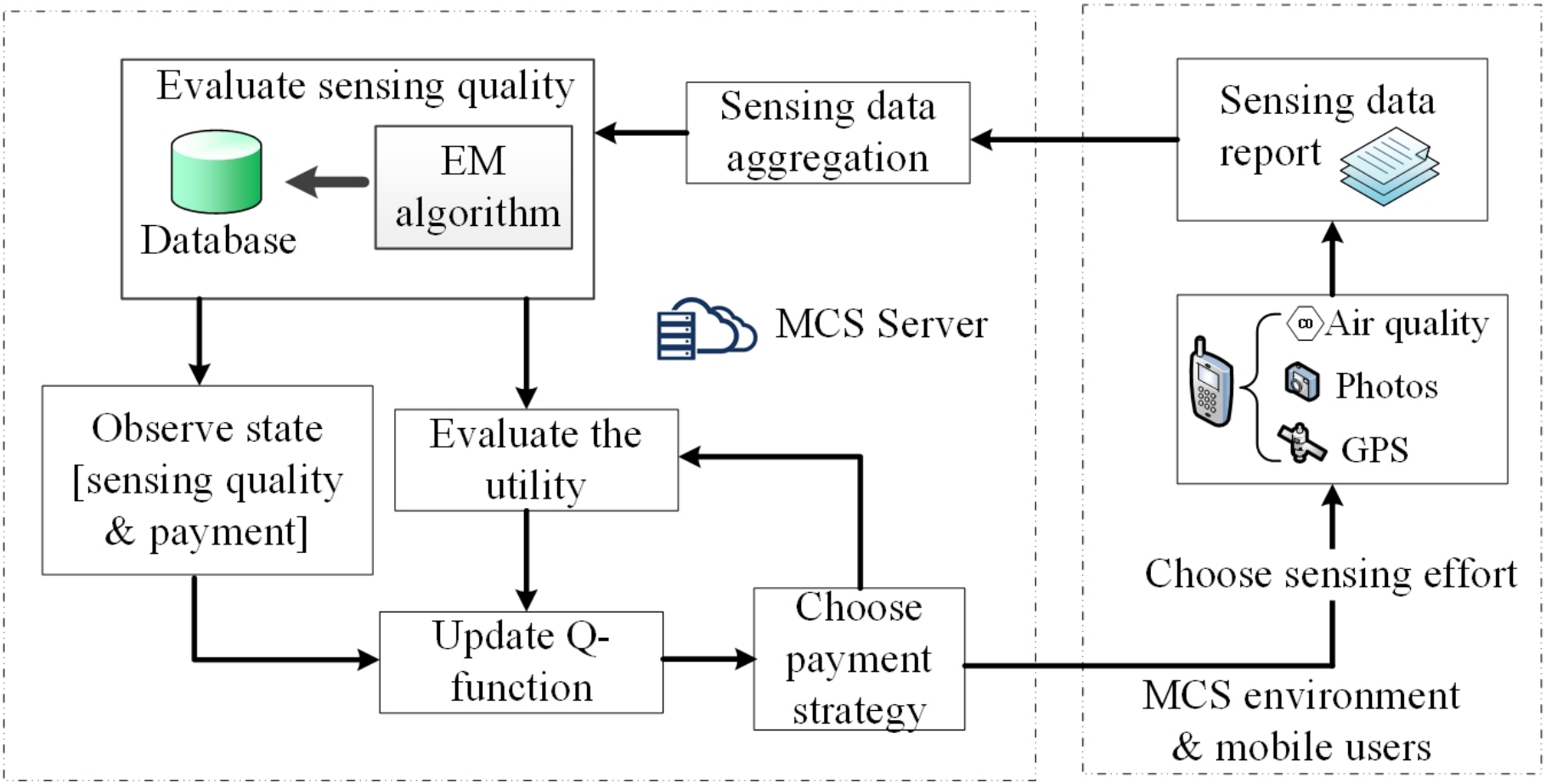
- MCS服务器的支付策略影响感知用户的感知策略
  - MCS过感知vs欠感知
  - 影响到服务器的未来效益
- 移动群智感知服务器的付费过程为马尔科夫决策过程（MDP）
- 考虑感知质量的评估错误，采用强化学习构造MCS激励机制
  - 质量评估错误会抑制合法用户的感知行为
  - 多种原因引发的感知质量下降，模型复杂

基于强化学习的移动群智感知

- 服务器采用强化学习技术，通过试错法获取最佳的定价策略
  - 无需预知虚假感知攻击和移动用户的感知模型
  - 适应动态网络拓扑结构的变化
  - 在有限状态的马尔科夫决策过程中实现最佳动态规划

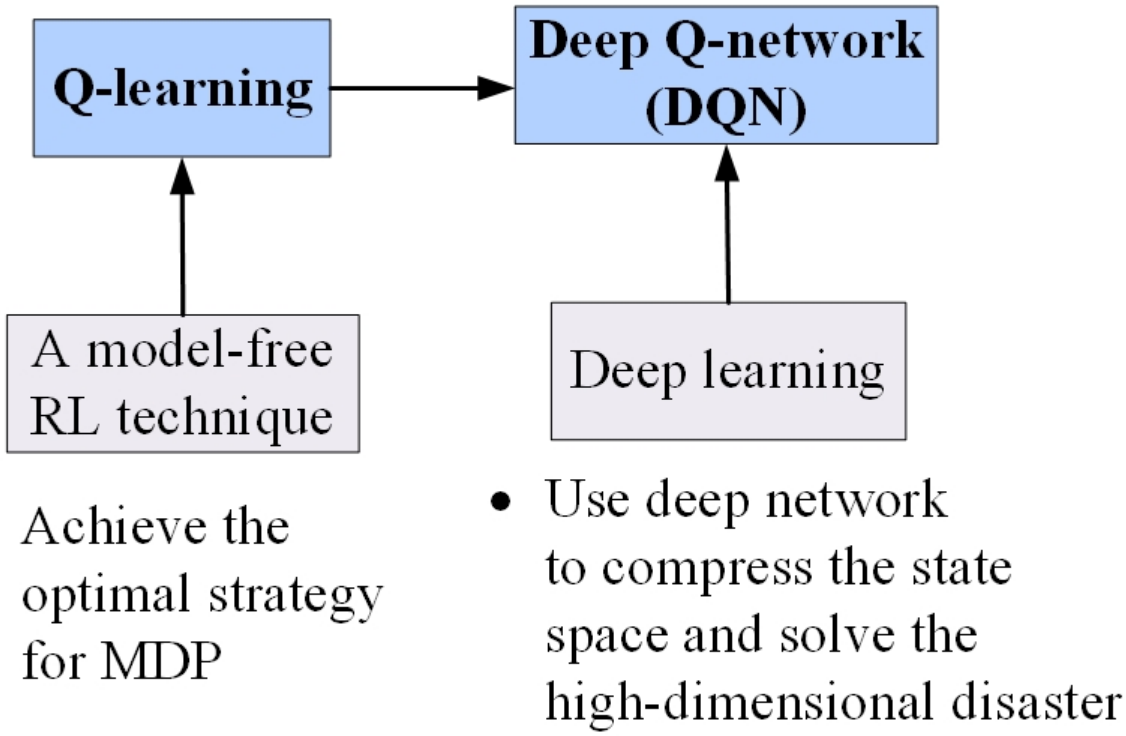


基于Q-learning的MCS系统

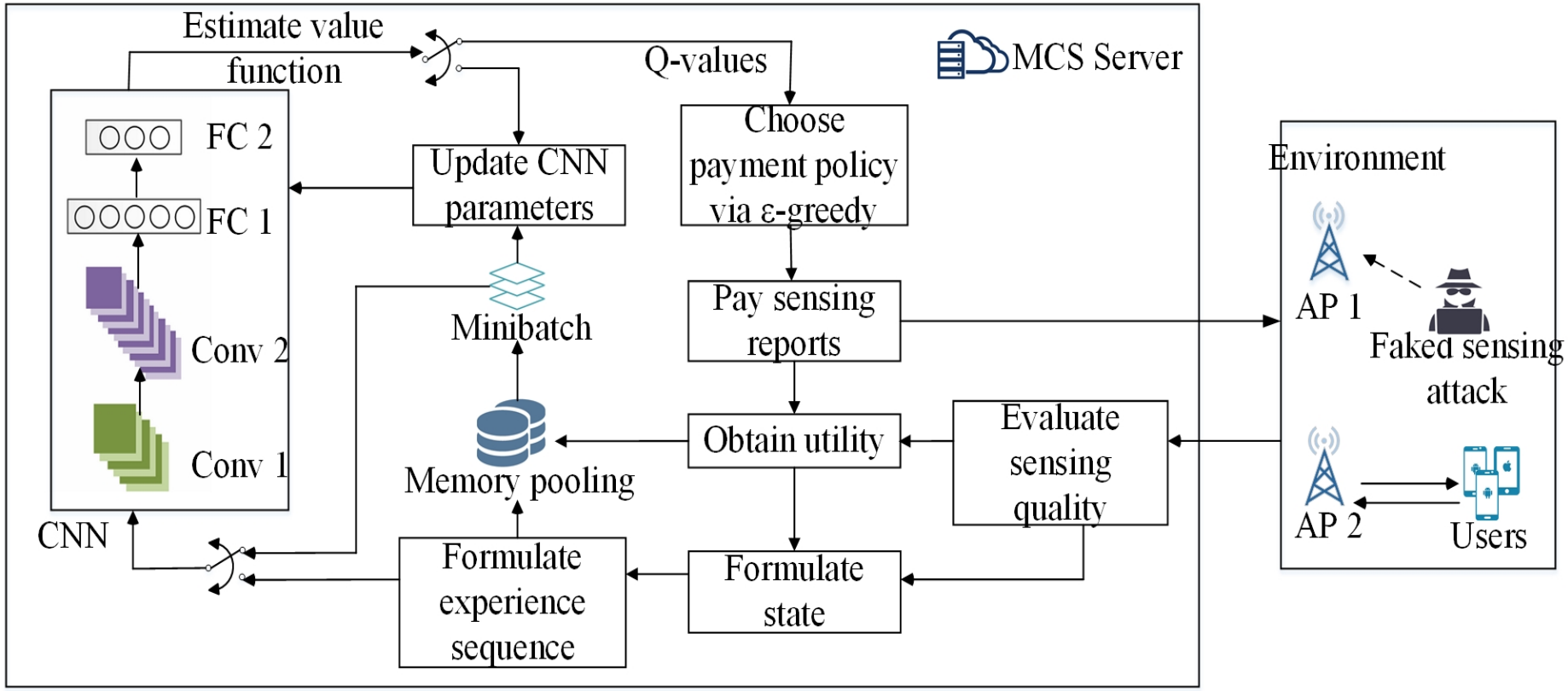


深度强化学习

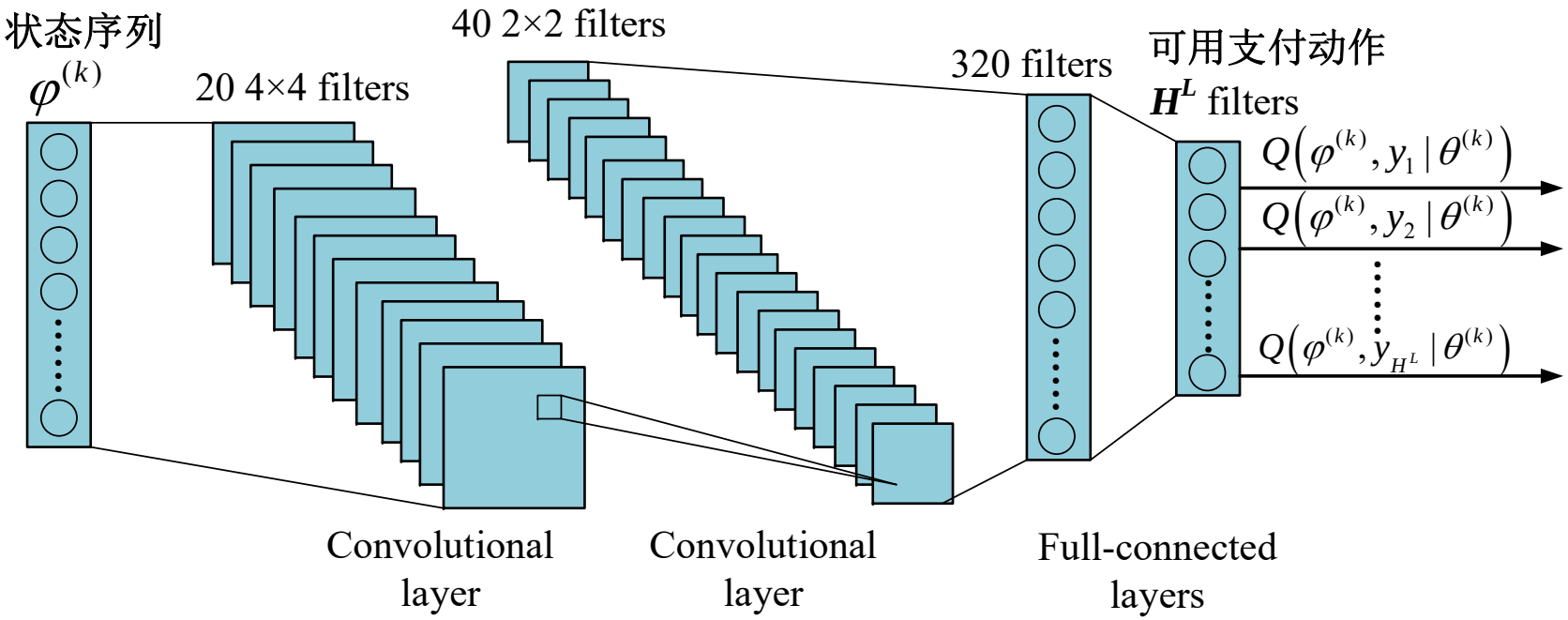
- 强化学习结合深入学习技术解决复杂决策任务
  - 快速提取数据特征，掌握动态环境的变化规律
  - 处理高维度状态信息，压缩状态空间



基于DQN的MCS系统



CNN模块



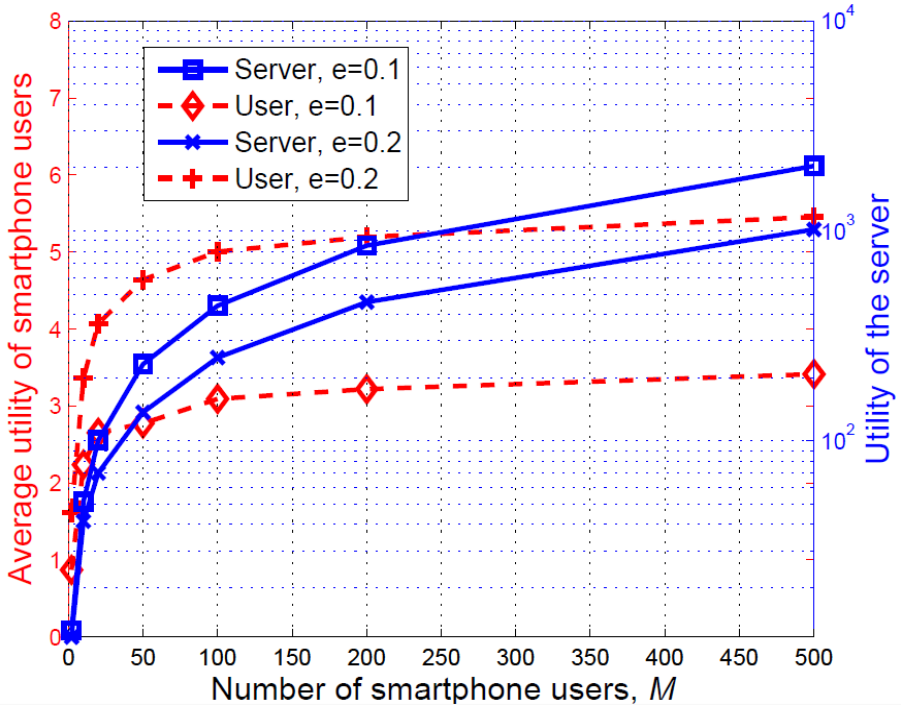
Layer	Input	Filter size	Stride	Number of filters	Activation	Output
Conv 1	8x8	4x4	1	20	ReLU	5x5x20
Conv 2	5x5x20	2x2	1	40	ReLU	4x4x40
FC 1	640			320	ReLU	320
FC 2	320			H^L	ReLU	H^L

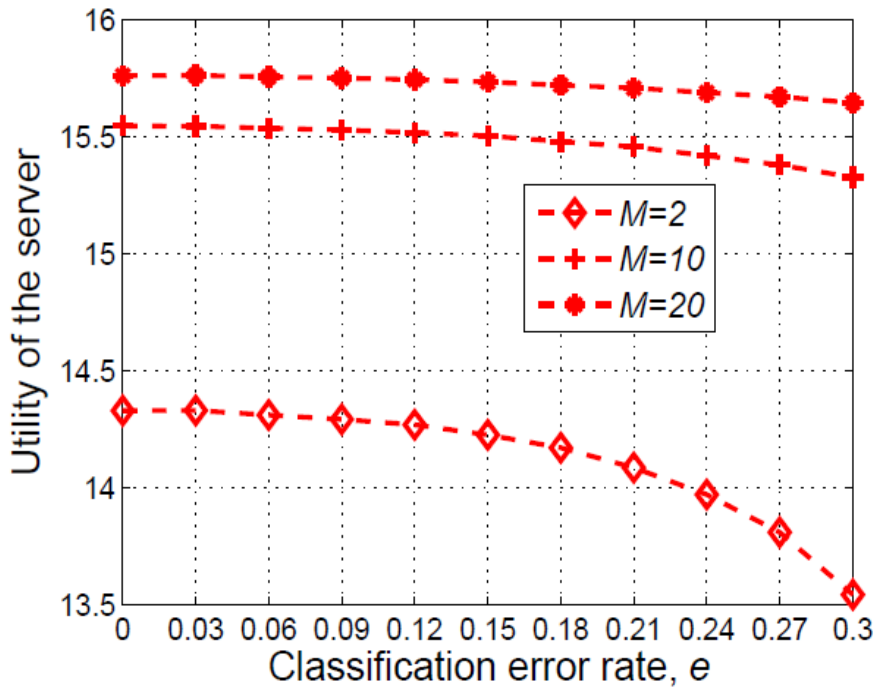
群智感知博弈均衡策略

- 建立移动群智感知Stackelberg博弈模型
  - 考虑用户感知损耗、贡献度因子、MCS服务器质量评估错误率
  - MCS服务器为lead,移动用户为follower
- 揭示各种条件下的MCS Stackelberg均衡:
  - 无法单方面地改变策略提高效益
  - 提供强化学习策略的MCS系统的性能界

L. Xiao, J. Liu, Q. Li, and H. Vincent Poor, "Secure mobile crowdsensing game," IEEE Int'l Conf. Commun., 2015.

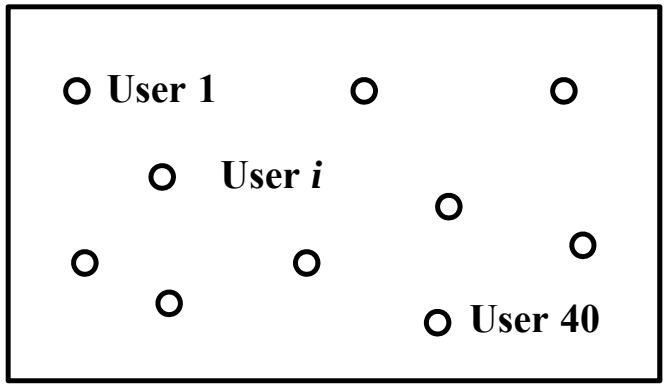
群智感知博弈Stackelberg均衡



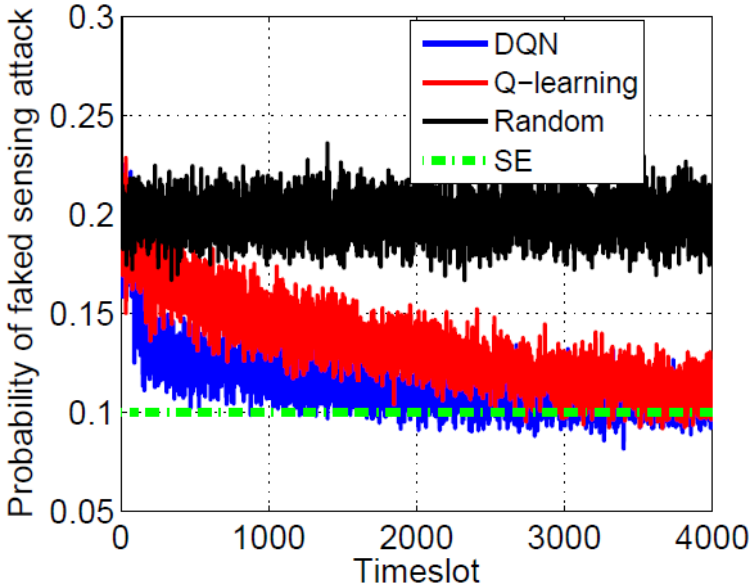


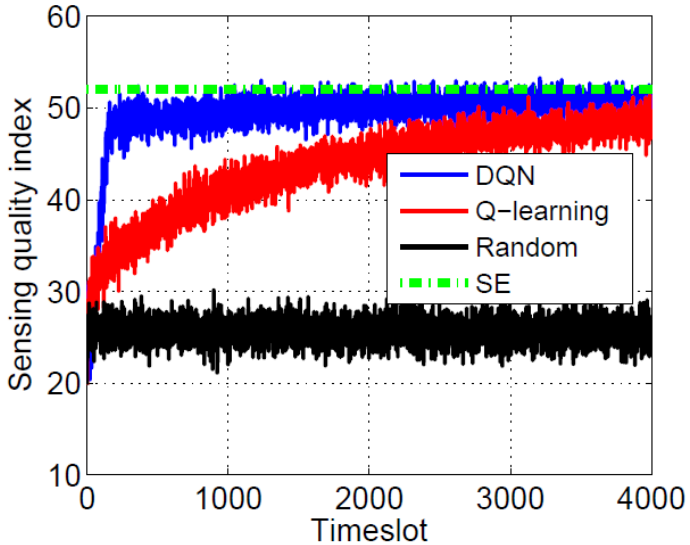
- ✓ 感知质量2级评估
- ✓ 用户贡献因子：$\lambda_j \sim U(0.5,1)$
- ✓ M个用户，感知错误率e

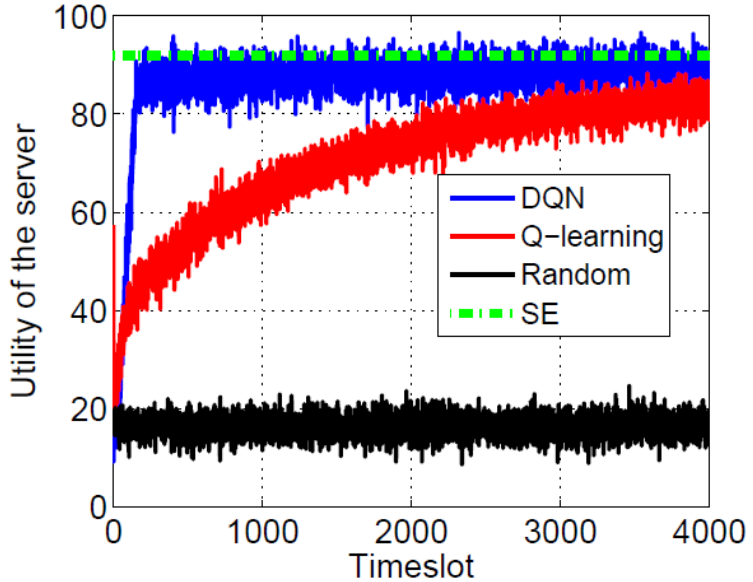
静态网络的MCS性能



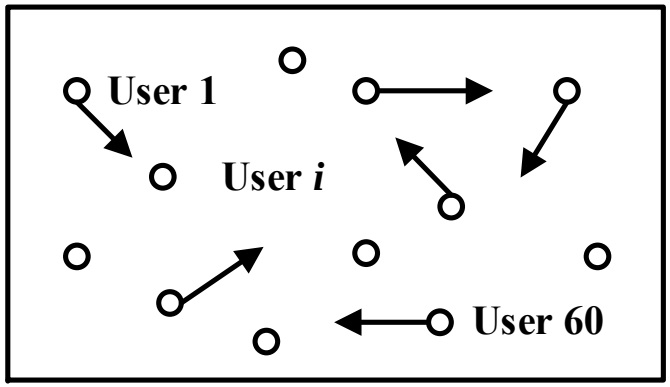
感知质量2级评估

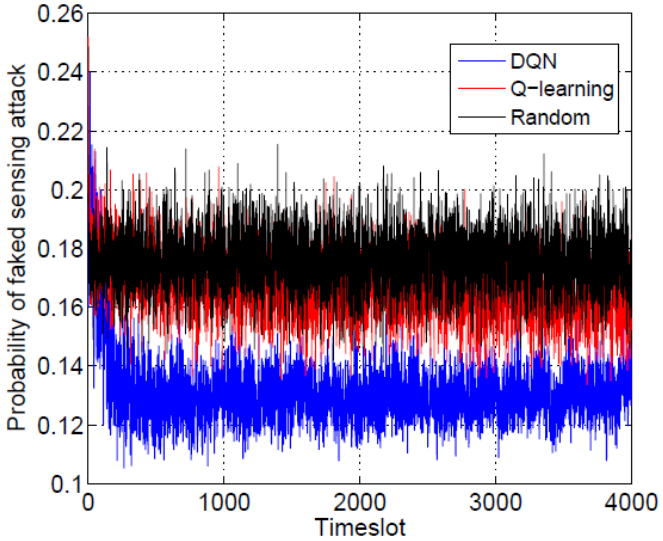




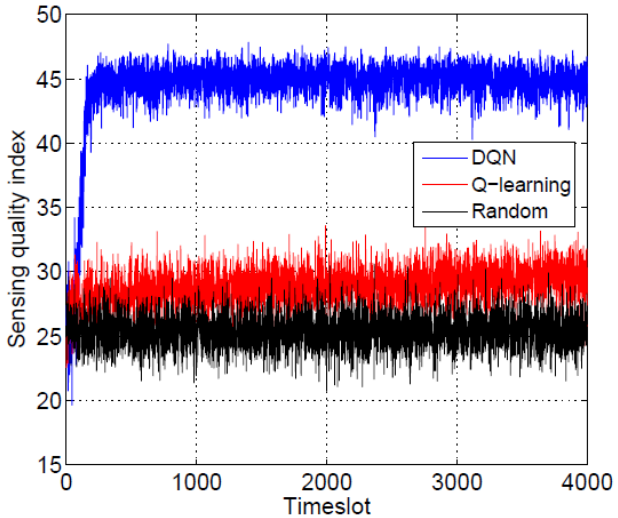


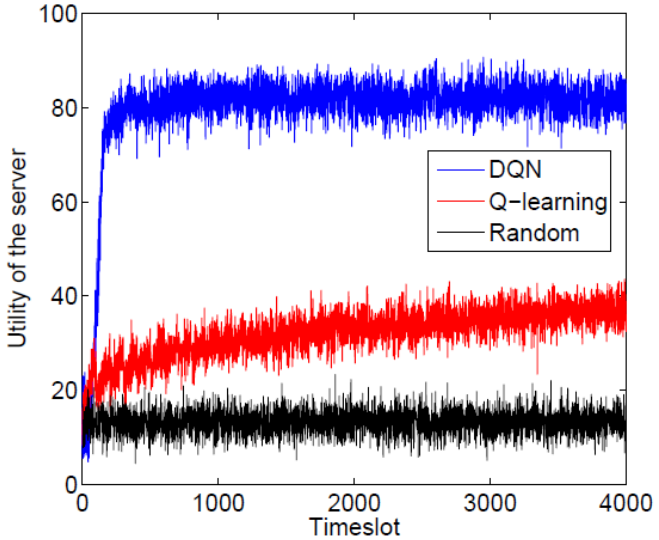
大规模移动网络下MCS系统性能





感知质量3级评估





L. Xiao, Y. Li, G. Han, H. Dai and H. Vincent Poor, "A Secure Mobile Crowdsensing Game with Deep Reinforcement Learning," IEEE Trans. Information Forensics & Security, 2018.

总结&未来工作

- 提出基于强化学习的移动群智感知定价机制，不需要预知移动用户的感知模型，抵御虚假感知攻击
- 采用深度强化学习算法，提高MCS在大规模移动网络中的感知质量
- 揭示博弈Stackelberg均衡，给出MCS系统的性能界
- 未来工作
  - 引入迁移学习技术，减少初始阶段的随机探索，加快学习速度
  - 构造安全备份机制，提供容灾措施

Questions?

lxiao@xmu.edu.cn

CNN更新策略

- DQN经验回放策略， 存储经验样本，将状态转移、支付策略选择、获取效益
  - 利用mini-batch，随机抽取批量为B的历史经验样本进行权重更新
  - 构造历史样本的目标损失函数
  - 利用随机梯度下降算法更新CNN参数