LocIn: Inferring Semantic Location from Spatial Maps in Mixed Reality

Habiba Farrukh, Reham Aburas, Aniket Nare, Antonio Bianchi, and Z. Berkay Celik

Purdue University

USENIX Security Symposium 2023

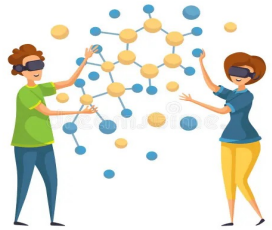
Mixed reality (MR)

MR devices allow users to view and interact with real and virtual content in their physical environment.

Retail

Entertainment

Healthcare



Education

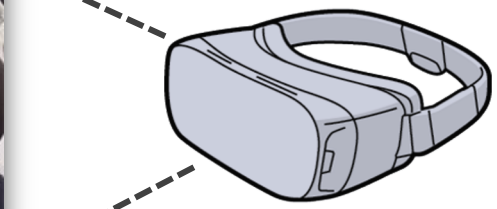
Military Training

Manufacturing

3D spatial maps in mixed reality

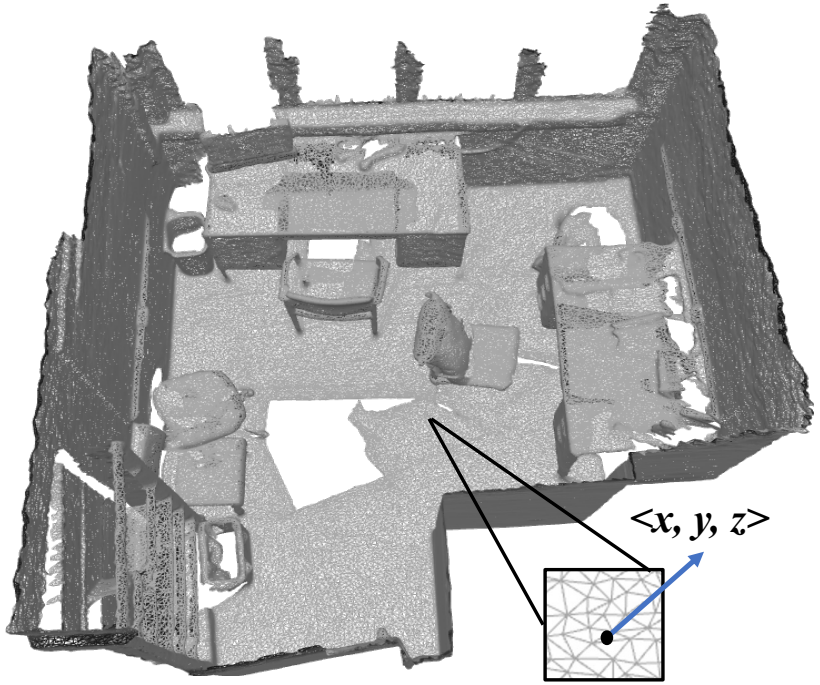
3D spatial maps - measure the distance between the device and points in the real environment.

User's environment



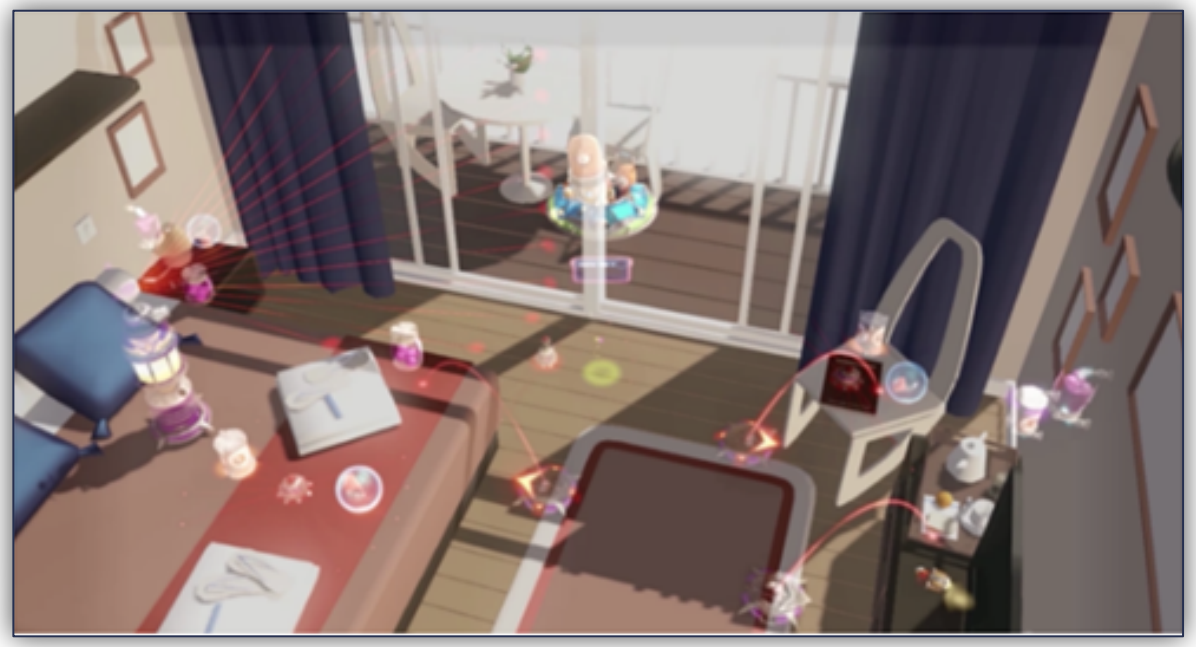
MR Device

Spatial map of user's environment

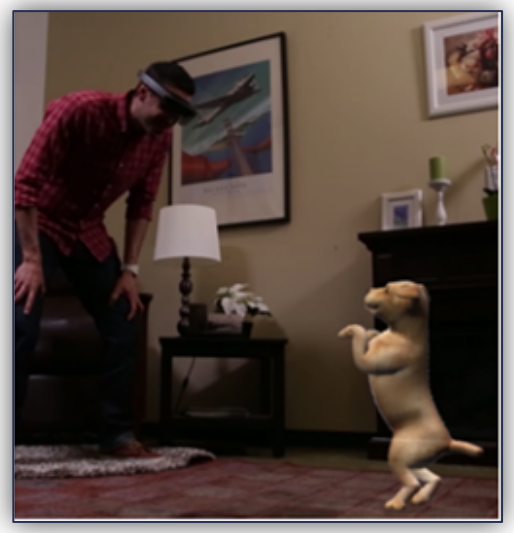


3D spatial maps in mixed reality

Object Placement



Occlusion Detection

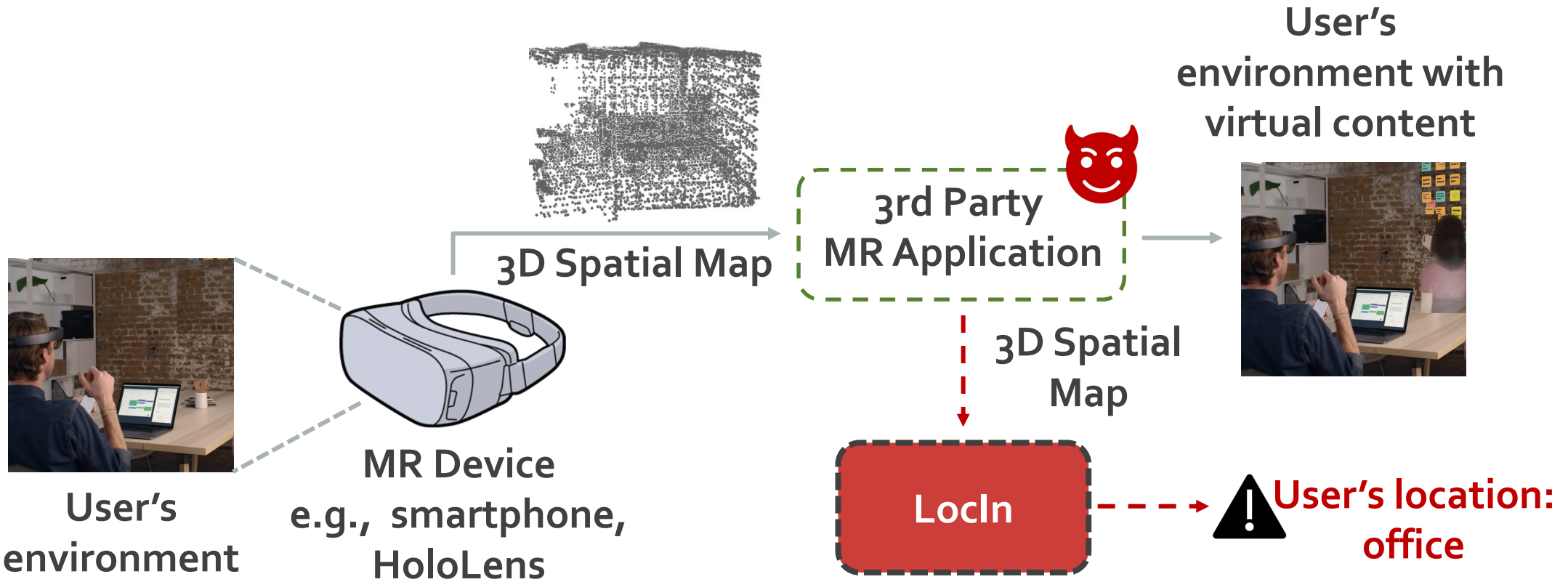


Navigation

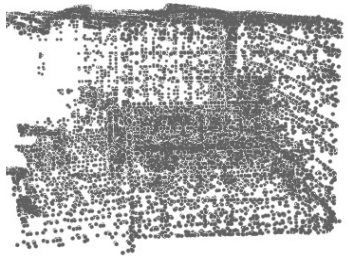
Source: https://learn.microsoft.com/en-us/windows/mixed-reality/design/spatial-mapping

LocIn: location inference from spatial maps in MR

Key insight: A malicious app can exploit the 3D spatial map of the user's environment to infer user's indoor locations i.e., semantic location.

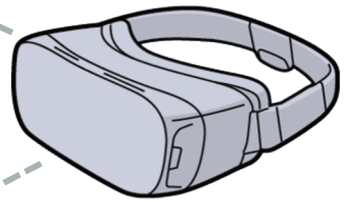


Threat model



3D Spatial Map

User's environment



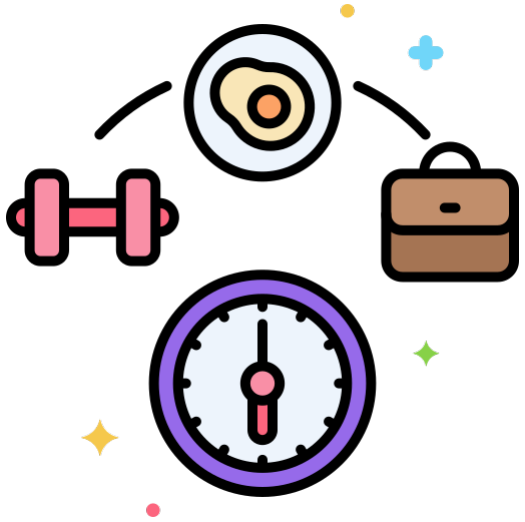
MR Device

3rd Party MR Application

- Aware of device type e.g., HoloLens/iPad
- Only requires camera access.
- Only uses point cloud → no color or normal vector information is needed.

Security and privacy implications

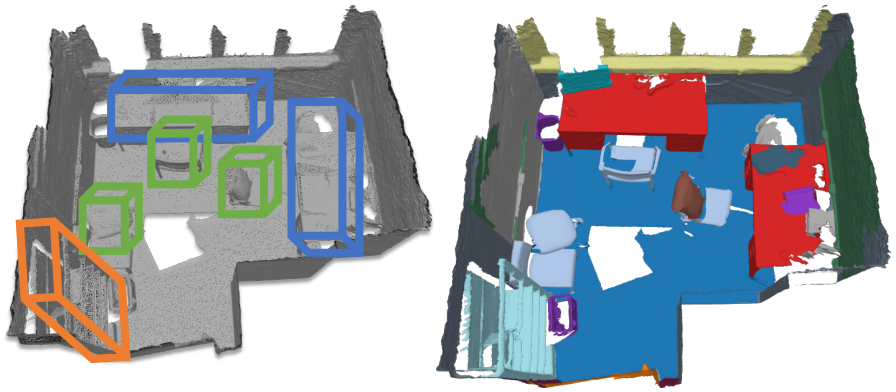
Robbery and Physical Attacks

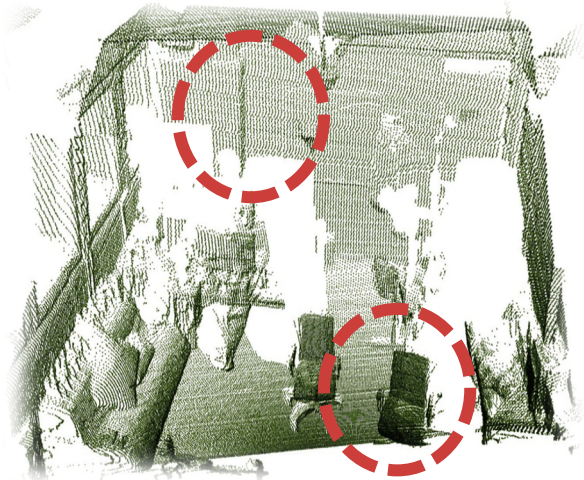


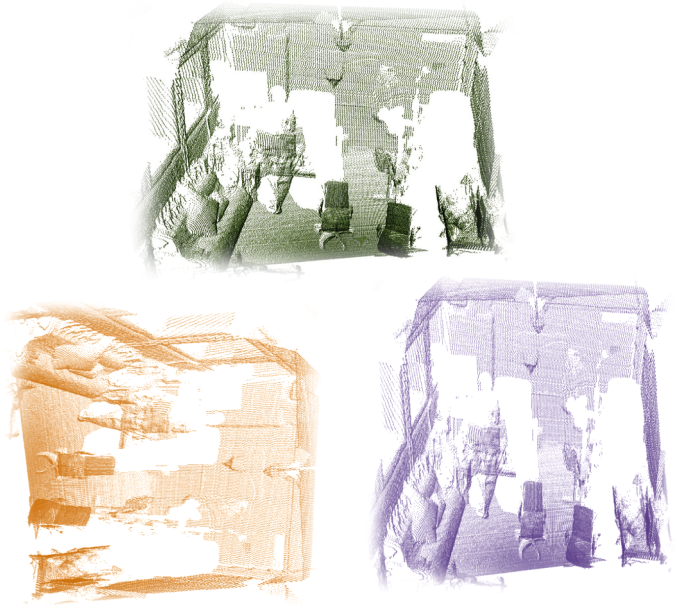
User Routine and Personalized Ads

Socio-economic and Family Status

Extracting location cues from spatial maps



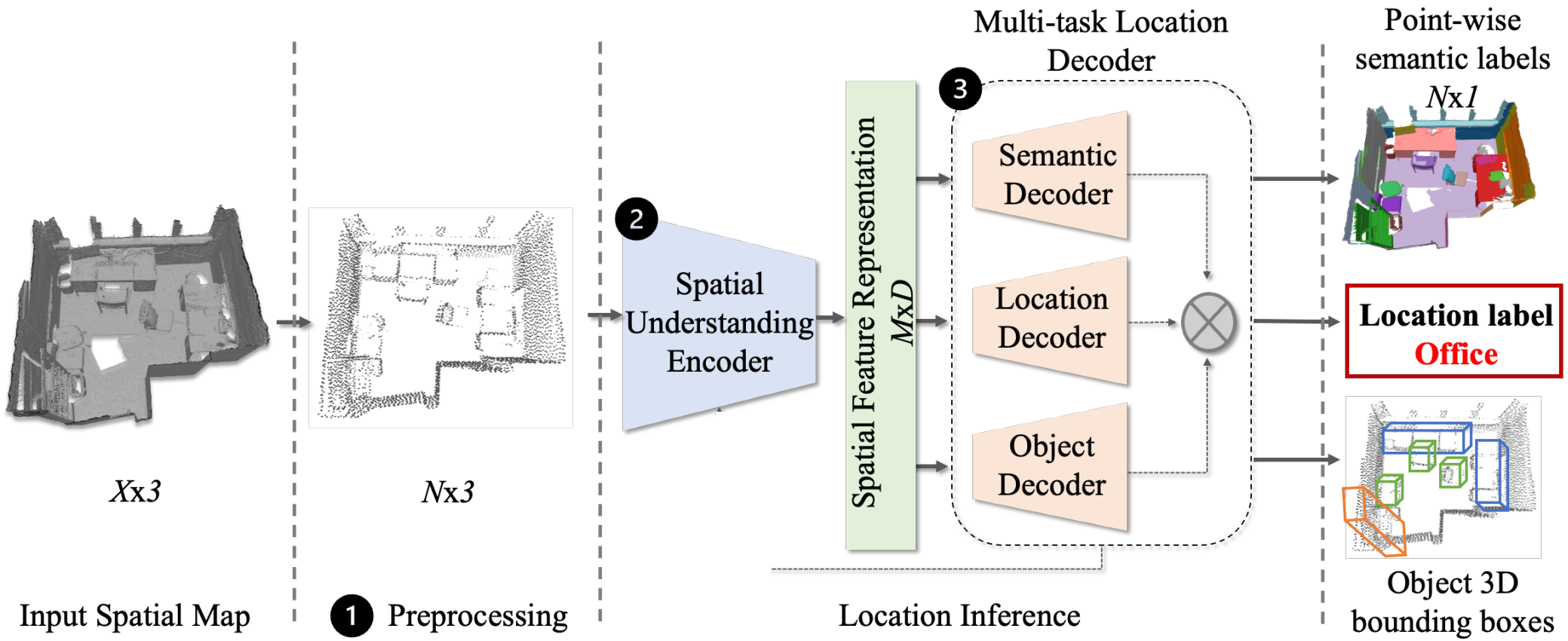




1. **Geometric and semantic properties**
2. **Non-uniform point density**
3. **Varying app usage**

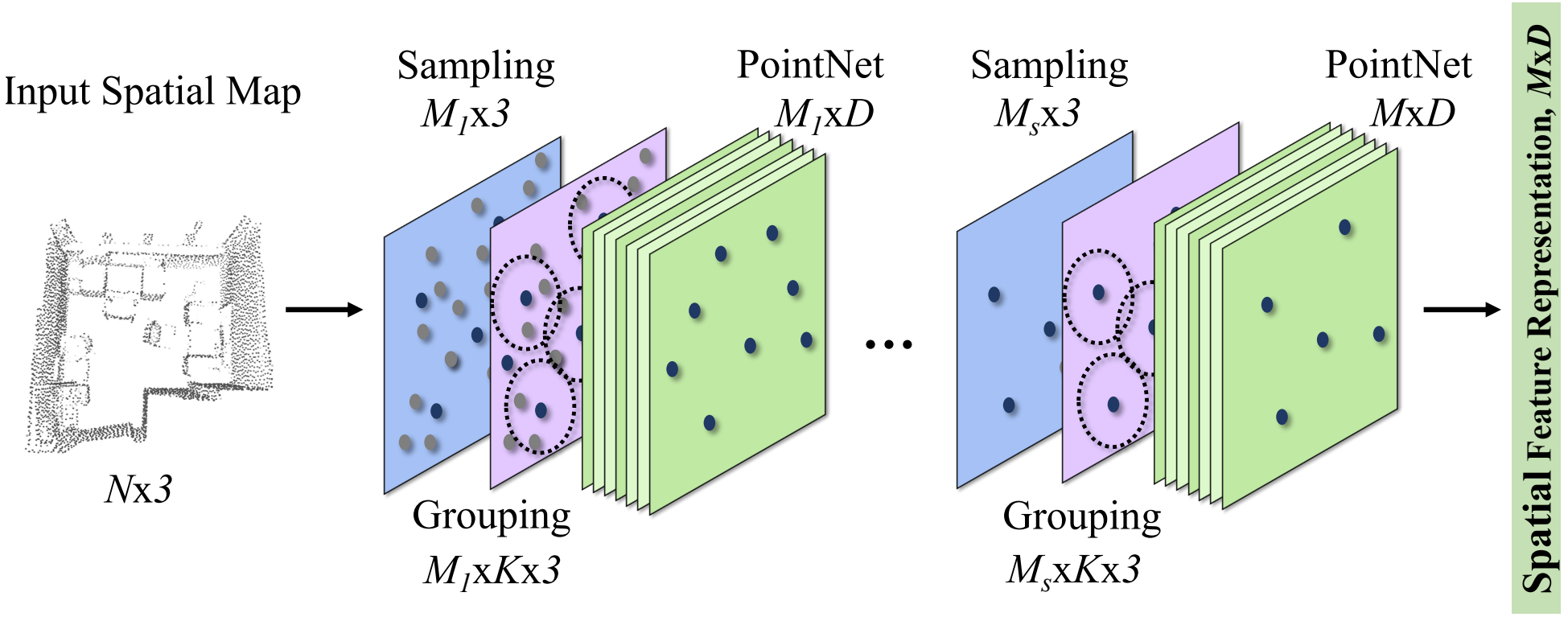
LocIn's overview

- We propose a novel location inference attack with end-to-end encoder-decoder architecture with a multi-task loss function.



LocIn's spatial encoder

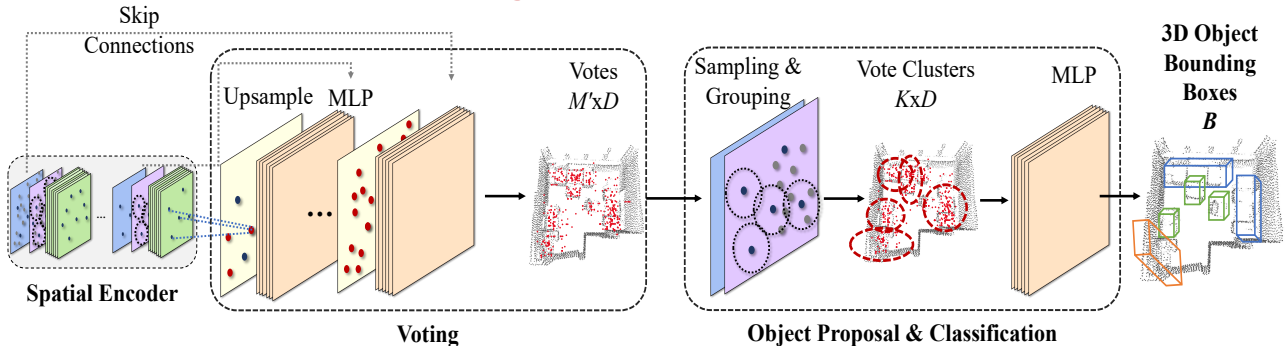
- We leverage a hierarchical encoder to capture the geometric and semantic properties of the user's environment.



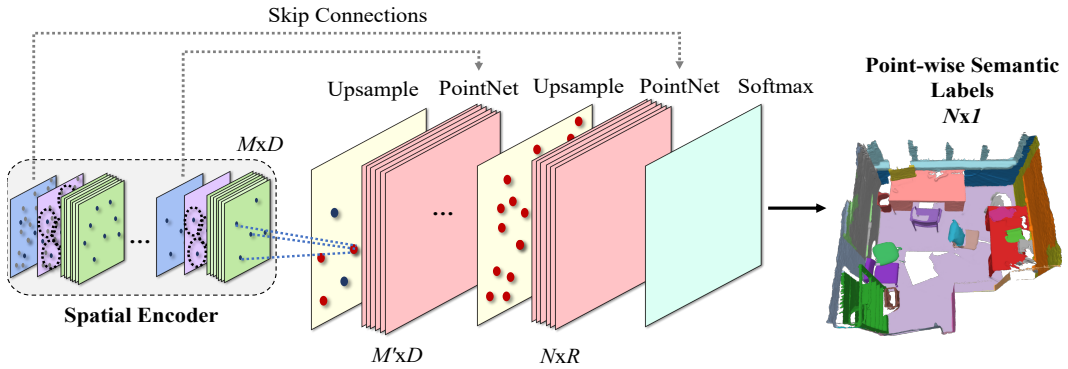
LocIn's multi-task decoder

- We propose a novel composite learning representation for location decoding by leveraging the multi-task learning paradigm.

Object Detection



Semantic Segmentation



$$f_s = \arg\min_{f \in \mathcal{F}_s} \frac{1}{n} \sum_{i=1}^{n} \left[\alpha L_{\text{loc}} + \beta L_{\text{obj}} + \gamma L_{\text{sem}} \right]$$

Location classification loss

Object detection loss

Semantic segmentation loss

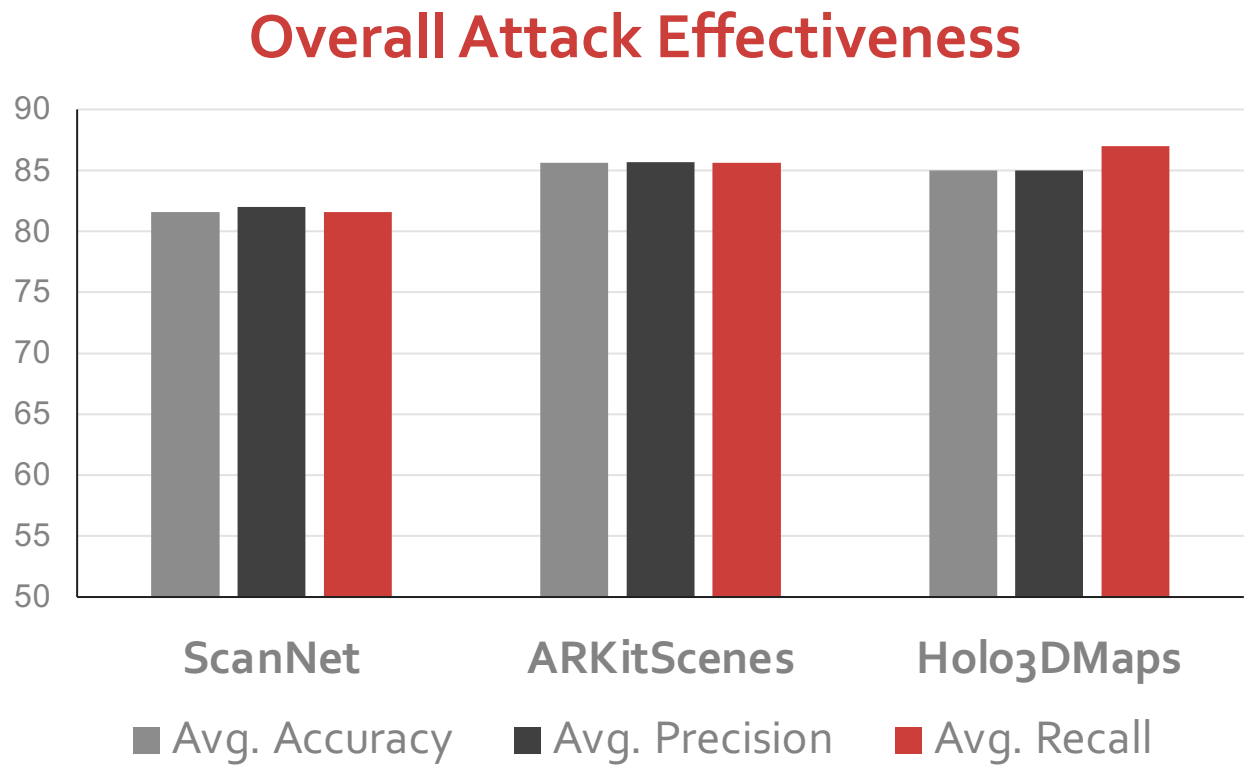
LocIn's evaluation

- Three datasets collected from different spatial sensors.

Dataset	MR Device	# of Location Classes	# of Object Classes	# of Spatial Maps
ScanNet	iPad Air2 with depth sensor	13	18	1513
ARKitScenes	iPad Pro with LiDAR scanner	9	17	5030
Holo3DMaps	HoloLens 2 with depth sensor	5	8	20

LocIn's effectiveness

- Evaluated on spatial map datasets collected from three popular MR devices.

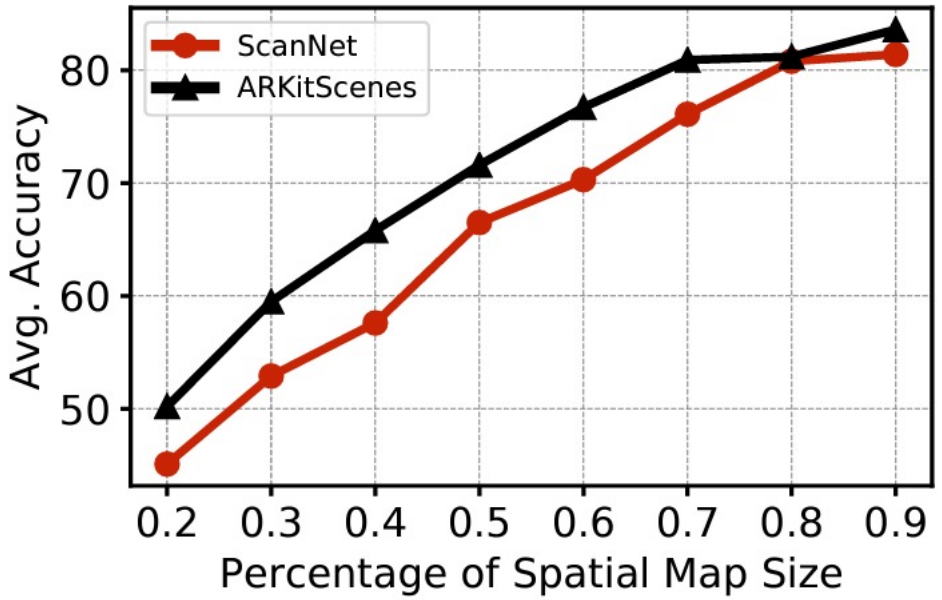


Infers location with >81% accuracy

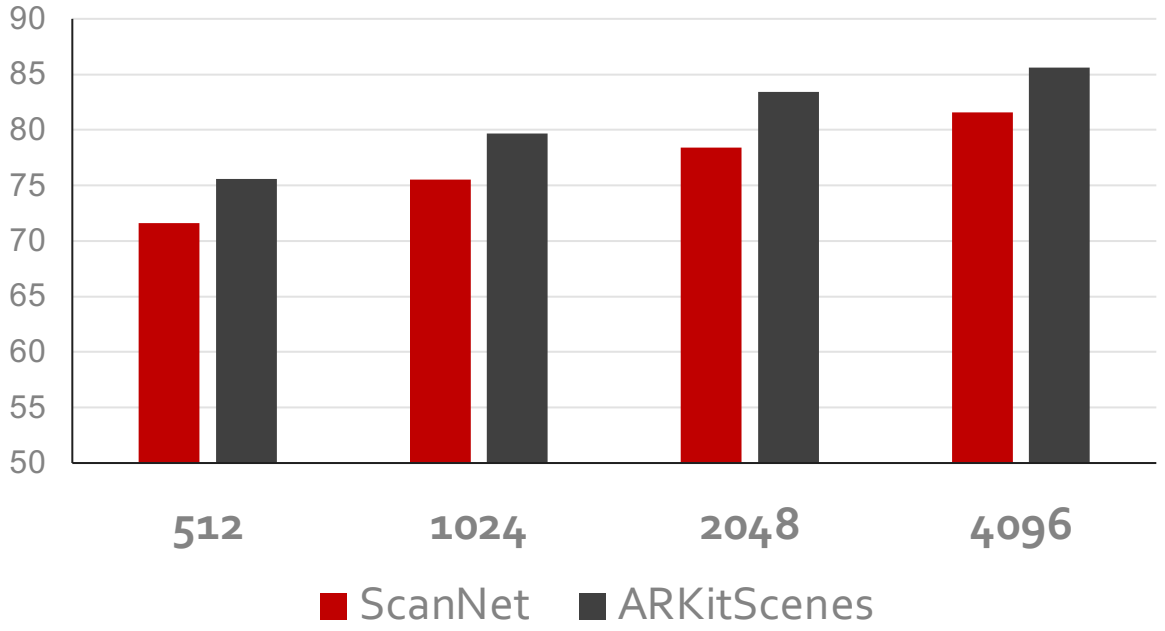
LocIn's robustness

- We demonstrate LocIn's robustness against varying spatial map size and sparsity.

Varying spatial map size

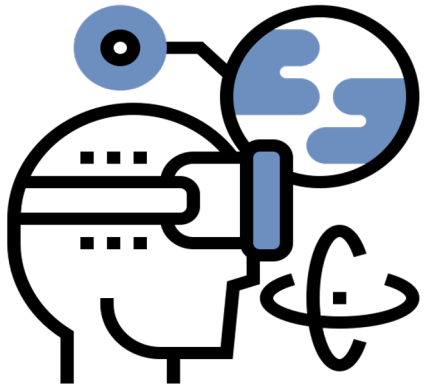


Varying spatial map sparsity



Conclusion

LocIn is a new location inference attack on mixed reality (MR) devices via 3D spatial data.



- Study attack effectiveness in complex environments.
- Investigate countermeasures with privacy-utility tradeoff.
- Explore users' perception of spatial maps' security & privacy.

Thank you! Questions?

hfarrukh@purdue.edu