ACM CCS 2024

RISiren: Wireless Sensing System Attacks via Metasurface

Chenghan Jiang[1], Jinjiang Yang[1], Xinyi Li[2]

Qi Li[2 4] , Xinyu Zhang[3], Ju Ren[2 4*]

Ubiquitous wireless sensing

Smart home

Wireless AI interprets omnipresent RF signals for pervasive sensing.
No Cameras, No Wearables, No Sensors
Wireless
Contactless
Sensorless
Rx
Tx

Health care

Intrusion detection

Ubiquitous wireless sensing

Smart home

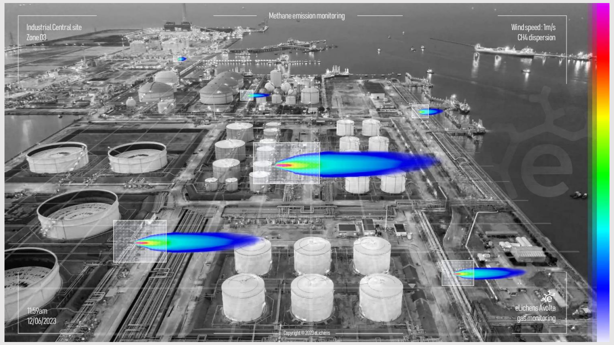
Wireless AI interprets omnipresent RF signals for pervasive sensing.
No Cameras, No Wearables, No Sensors
Wireless
Contactless
Sensorless
Tx
Rx

Health care

Intrusion detection

Intelligent transportation

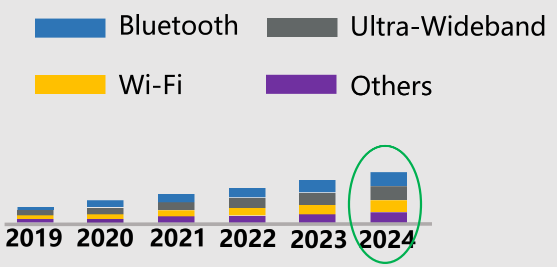
Security authentication



Indicator monitoring

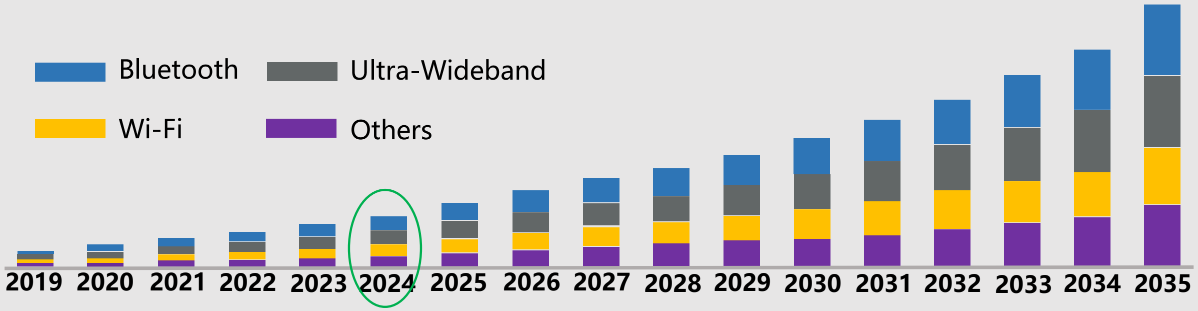
Ubiquitous wireless sensing

USD 18.92 Billion Value in 2024



Data from the statistical report of Roots Analysis organization: Wireless Sensor Market, Till 2035

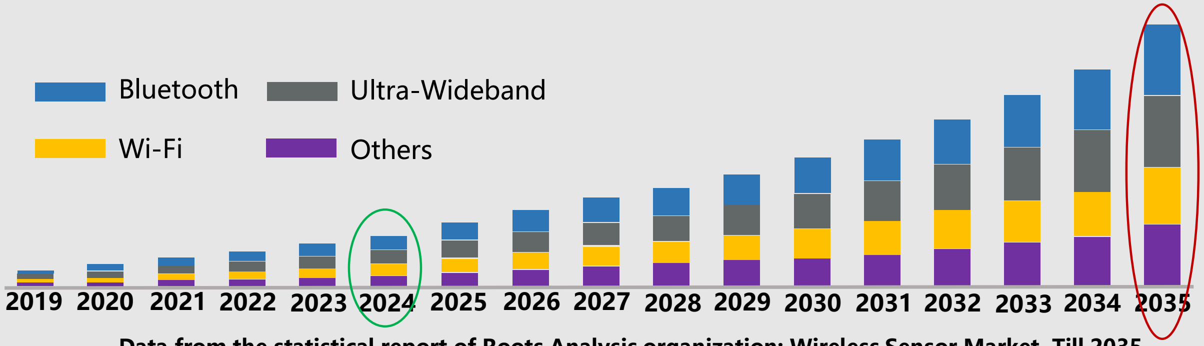
Ubiquitous wireless sensing



Data from the statistical report of Roots Analysis organization: Wireless Sensor Market, Till 2035

USD 18.92 Billion Value in 2024

Ubiquitous wireless sensing

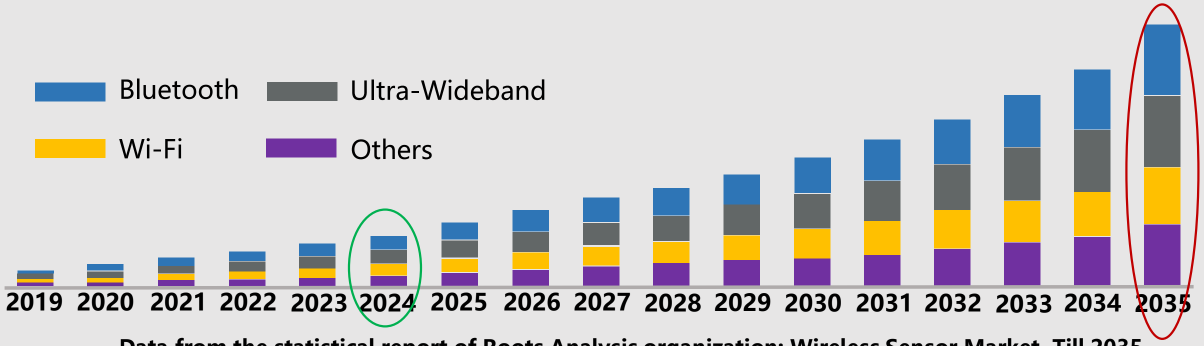


Data from the statistical report of Roots Analysis organization: Wireless Sensor Market, Till 2035

USD 18.92 Billion Value in 2024

USD 159.03 Billion Value in 2035

Ubiquitous wireless sensing



Data from the statistical report of Roots Analysis organization: Wireless Sensor Market, Till 2035

USD 18.92 Billion Value in 2024

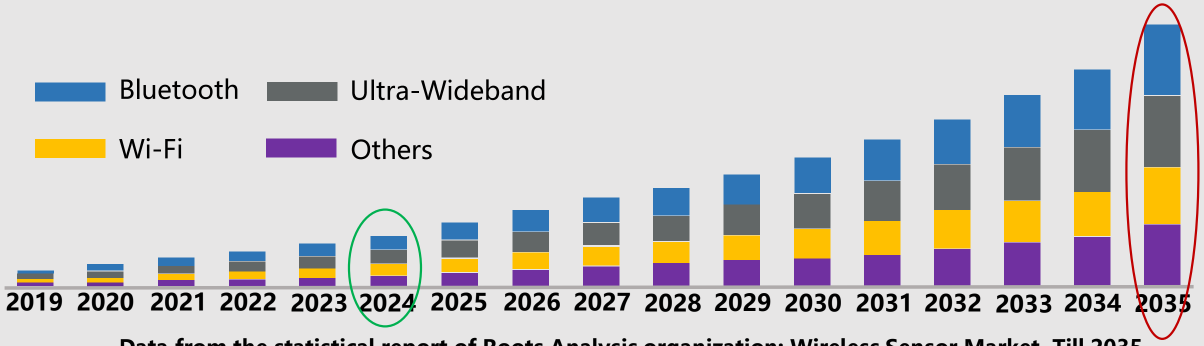
USD 159.03 Billion Value in 2035

21.35% CAGR 2024-2035

"Zoe Fall is a manifestation of our mission to help the elderly maintain their independence. Our innovative Wi-Fi-based fall-detection solution respects privacy and offers peace of mind for millions of senior citizens."

Thomas Saphir, Zoe Care's CEO

Ubiquitous wireless sensing



Data from the statistical report of Roots Analysis organization: Wireless Sensor Market, Till 2035

USD 18.92 Billion Value in 2024

USD 159.03 Billion Value in 2035

21.35% CAGR 2024-2035

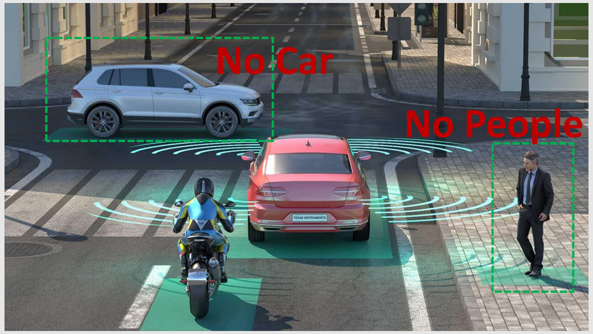
"Zoe Fall is a manifestation of our mission to help the elderly maintain their independence. Our innovative Wi-Fi-based fall-detection solution respects privacy and offers peace of mind for millions of senior citizens."

Thomas Saphir, Zoe Care's CEO

Can wireless sensing be fully reliable?

The deep penetration of **wireless sensing** has gradually exposed fatal problems
due to the **broadcast nature of wireless media**

The deep penetration of **wireless sensing** has gradually exposed fatal problems due to the **broadcast nature of wireless media**



Interfere intelligent driving[1]

Tamper voice assistant[2]

Deceit intrusion detection[3]

[1] [CCS'23] TileMask: A Passive-Reflection-based Attack against mmWave Radar Object Detection in Autonomous Driving

[2] [NDSS'24] Inaudible Adversarial Perturbation: Manipulating the Recognition of User Speech in Real-Time

[3] [Sensys'23] RIStealth: Practical and Covert Physical-Layer Attack against WiFi-based Intrusion Detection via Reconfigurable Intelligent Surface

Prior works limitation

Prior works limitation

High requirements for attackers

Prior works limitation

High requirements for attackers

Assume victim system framework can be known or learned[1]

[1] [Mobicom'22] Audio-domain Position-independent Backdoor Attack via Unnoticeable Triggers

Prior works limitation

High requirements for attackers

Assume victim system framework can be known or learned[1]

High detectability for attackers

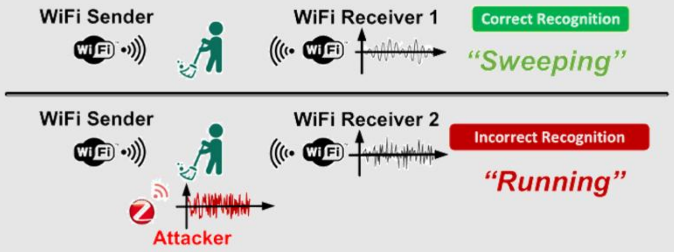
[1] [Mobicom’22] Audio-domain Position-independent Backdoor Attack via Unnoticeable Triggers

Prior works limitation

High requirements for attackers

Assume victim system framework can be known or learned[1]

High detectability for attackers



Extra active devices to execute attacks[2]

[1] [Mobicom'22] Audio-domain Position-independent Backdoor Attack via Unnoticeable Triggers

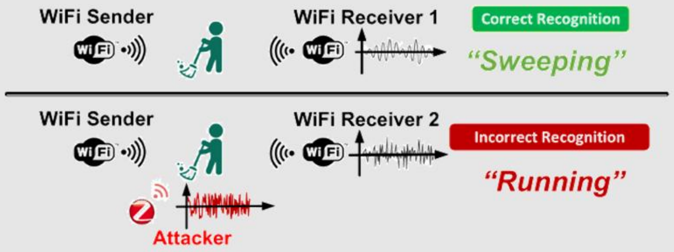
[2] [IEEE TDSC] IS-WARS: Intelligent and Stealthy Adversarial Attack to Wi-Fi-Based Human Activity Recognition Systems

Prior works limitation

High requirements for attackers

Assume victim system framework can be known or learned[1]

High detectability for attackers



Extra active devices to execute attacks[2]

High cost and form factors

[1] [Mobicom'22] Audio-domain Position-independent Backdoor Attack via Unnoticeable Triggers

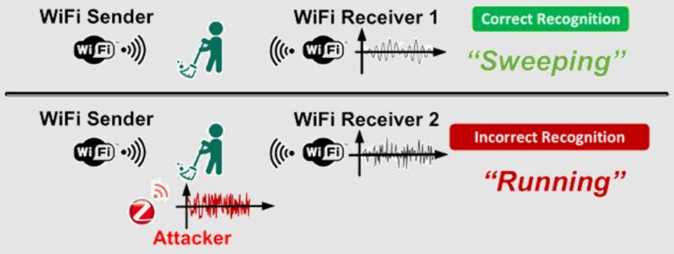
[2] [IEEE TDSC] IS-WARS: Intelligent and Stealthy Adversarial Attack to Wi-Fi-Based Human Activity Recognition Systems

Prior works limitation

High requirements for attackers

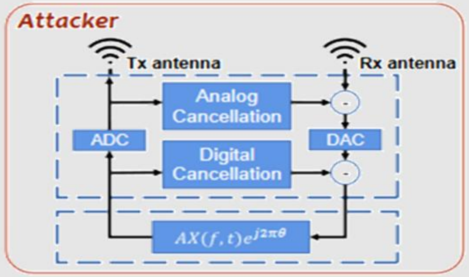
Assume victim system framework can be known or learned[1]

High detectability for attackers



Extra active devices to execute attacks[2]

High cost and form factors



Full-duplex devices to edit and transfer signal[3]

[1] [Mobicom'22] Audio-domain Position-independent Backdoor Attack via Unnoticeable Triggers

[2] [IEEE TDSC] IS-WARS: Intelligent and Stealthy Adversarial Attack to Wi-Fi-Based Human Activity Recognition Systems

[3] [Ubicomp'22] WiAdv: Practical and Robust Adversarial Attack against WiFi-based Gesture Recognition System

Is there a more threatening attack strategy that is stealthier and does not require the victim's prior knowledge?

Our work RISiren[1]

[1] "RISiren" derived from the sea-nymphs "Siren" who lured sailors to their death with a bewitching song in ancient Greek mythology

Our work RISiren[1]

[1] "RISiren" derived from the sea-nymphs "Siren" who lured sailors to their death with a bewitching song in ancient Greek mythology

Our work RISiren[1]

[1] “RISiren” derived from the sea-nymphs “Siren” who lured sailors to their death with a bewitching song in ancient Greek mythology

Our work RISiren[1]

Q1: How to attack invisibly without extra sources?

Q2: How to achieve a black-box attack?

[1] "RISiren" derived from the sea-nymphs "Siren" who lured sailors to their death with a bewitching song in ancient Greek mythology

Our work RISiren[1]

Q1: How to attack invisibly without extra sources?

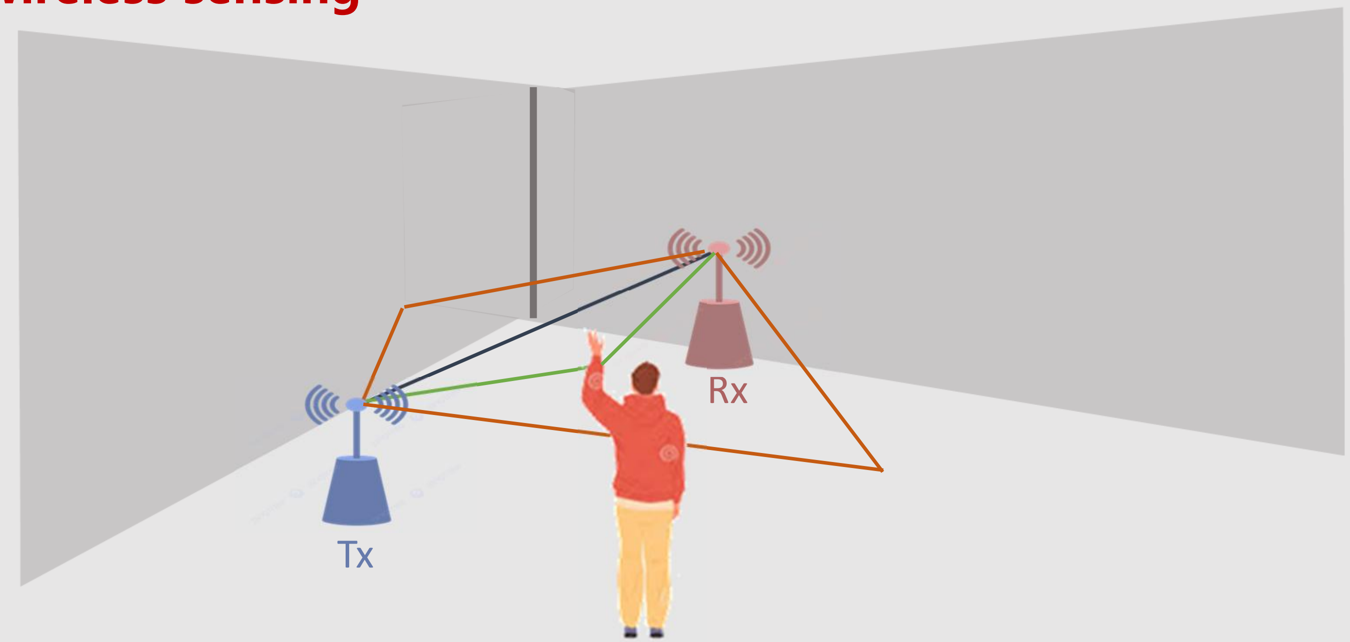
Q2: How to achieve a black-box attack?

[1] “RISiren” derived from the sea-nymphs “Siren” who lured sailors to their death with a bewitching song in ancient Greek mythology

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

Principle of wireless sensing

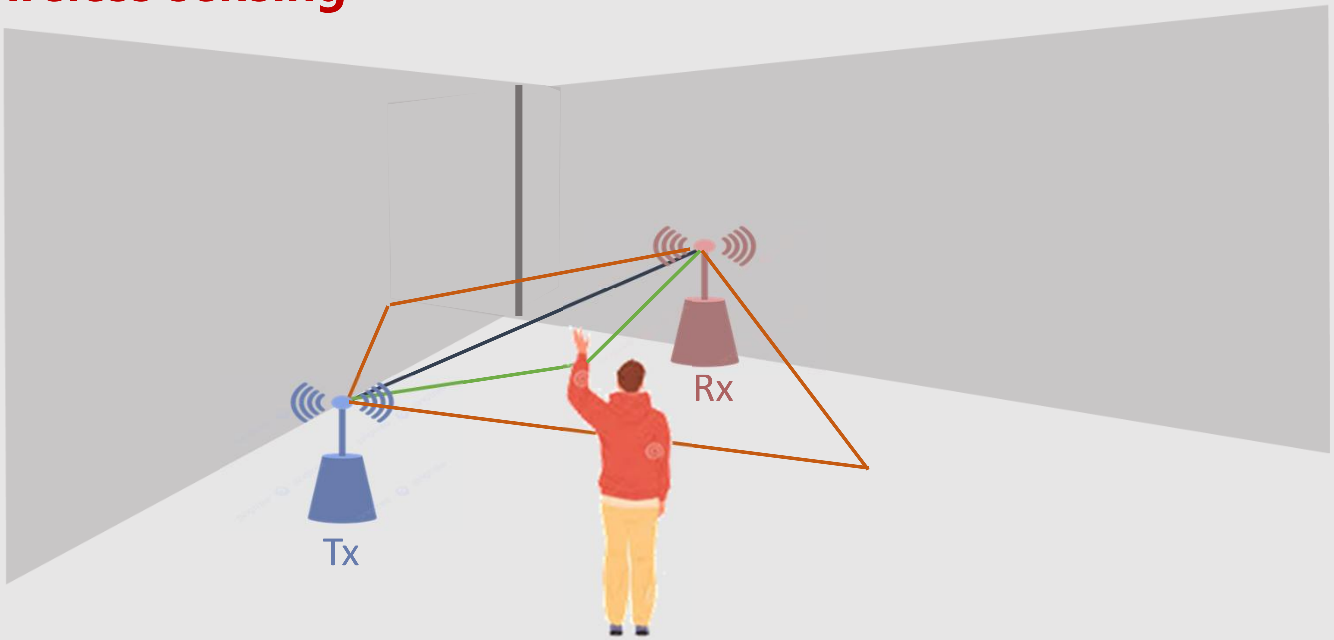


— Signal from the straight path — Signal from the multi-path — Signal from human activity

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

Principle of wireless sensing



— Signal from the straight path — Signal from the multi-path — Signal from human activity

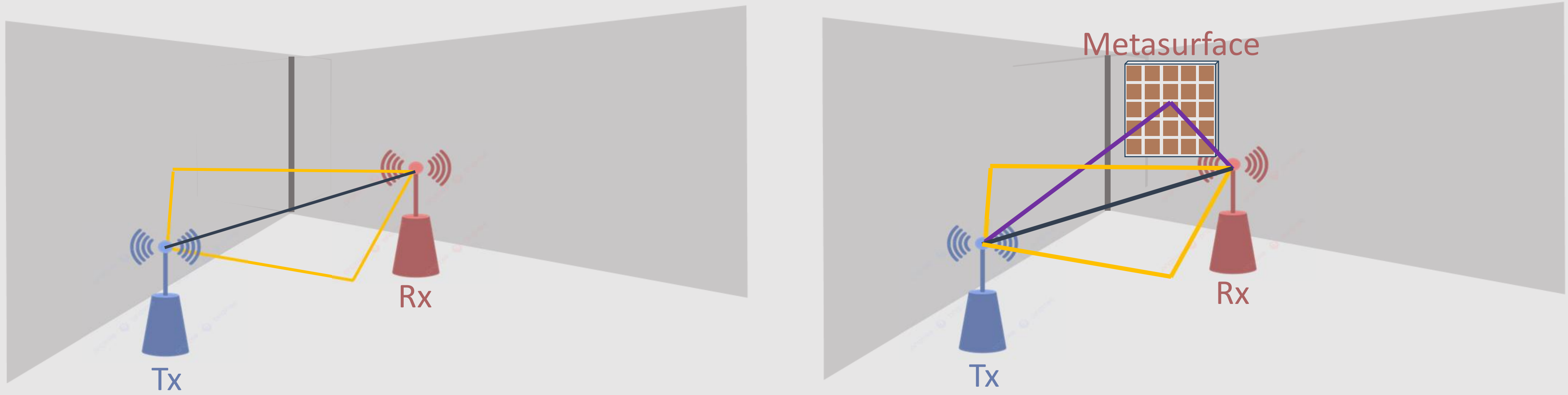
The results of wireless sensing will include reflected multi-path links

Insight of RISiren:
Can we generate a malicious multi-path to inject attack

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

Metasurface: **Reshape electromagnetic waves in space freely like a “mirror”**

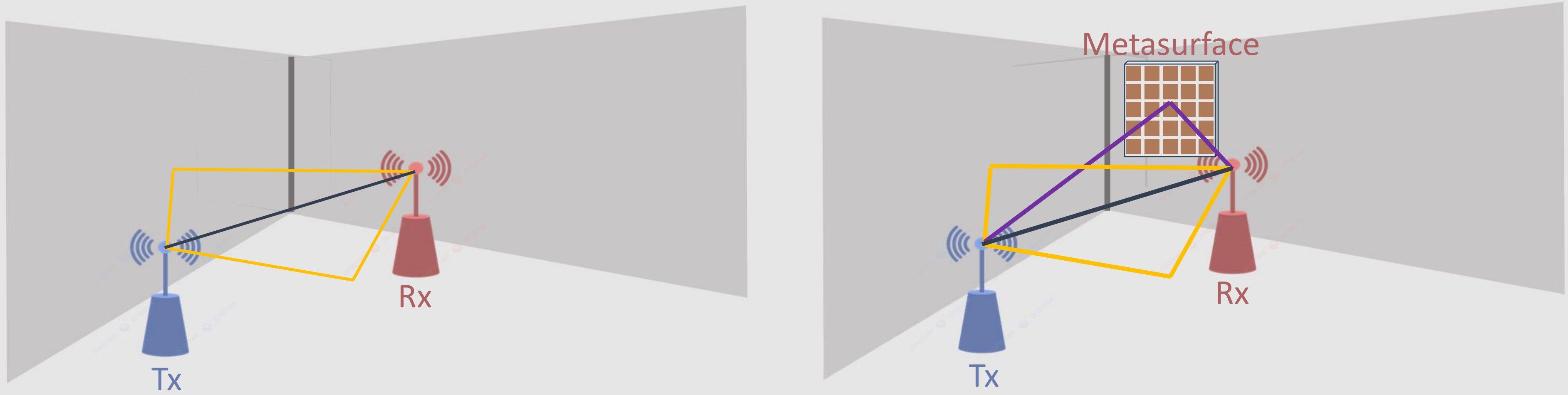


— Signal from the straight path — Signal from the multi-path — Signal from metasurface reflection

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

Metasurface: **Reshape electromagnetic waves in space freely like a “mirror”**



— Signal from the straight path — Signal from the multi-path — Signal from metasurface reflection

✦ The metasurface only **reshapes the signal in environment** instead of generating extra signal

RISiren - To address Q1

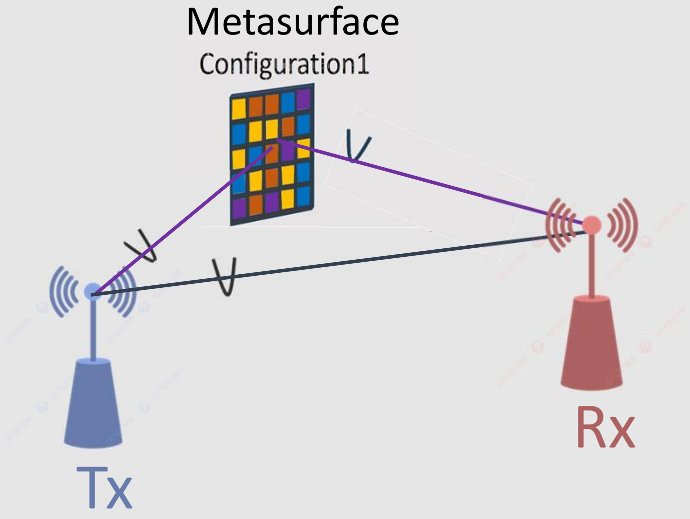
(Q1) How to attack invisibly without extra sources?

Key observation: Switching different metasurface configurations can generate time-variant interference to the wireless channel.

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

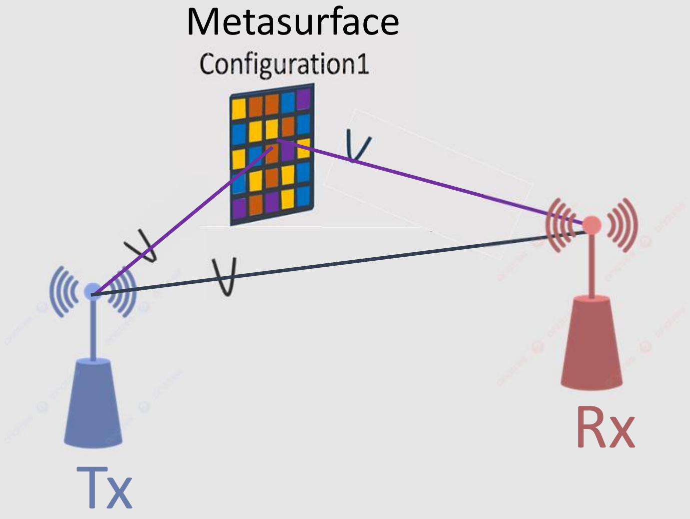
Key observation: Switching different metasurface configurations can generate time-variant interference to the wireless channel.

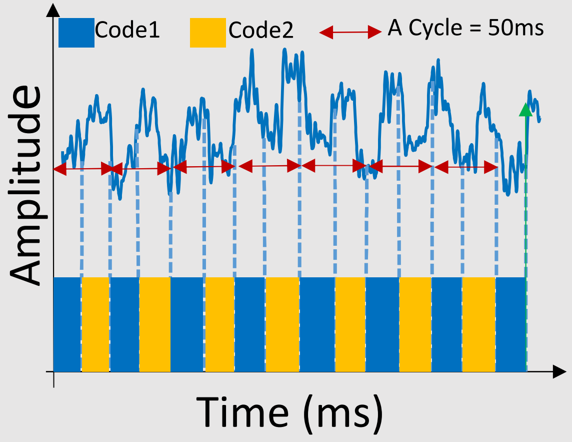


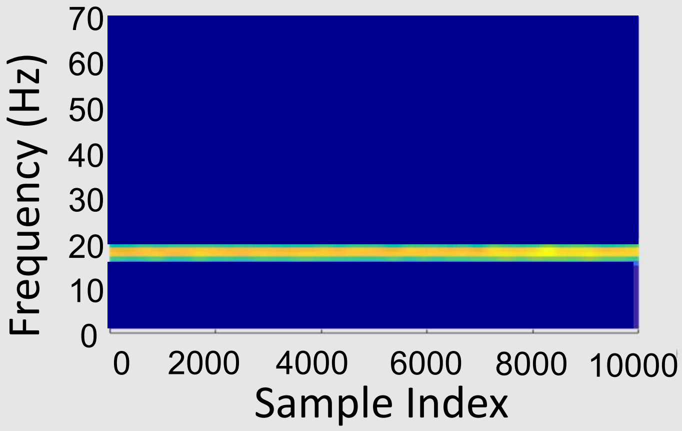
RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

Key observation: Switching different metasurface configurations can generate time-variant interference to the wireless channel.



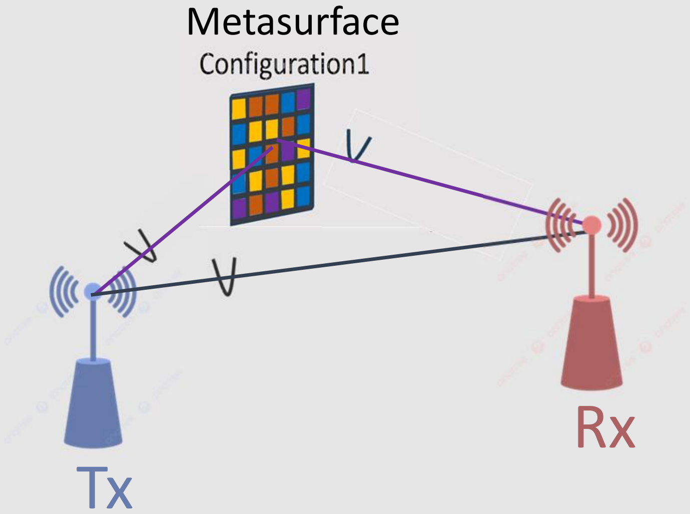


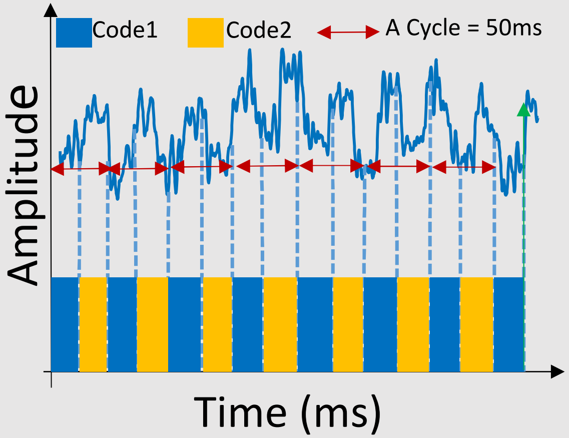


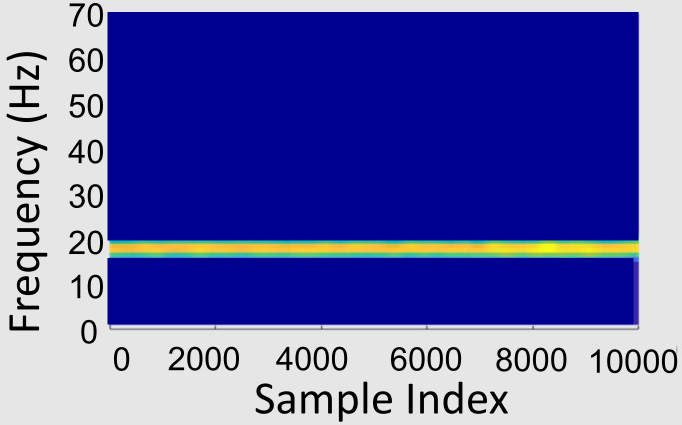
RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

Key observation: Switching different metasurface configurations can generate time-variant interference to the wireless channel.





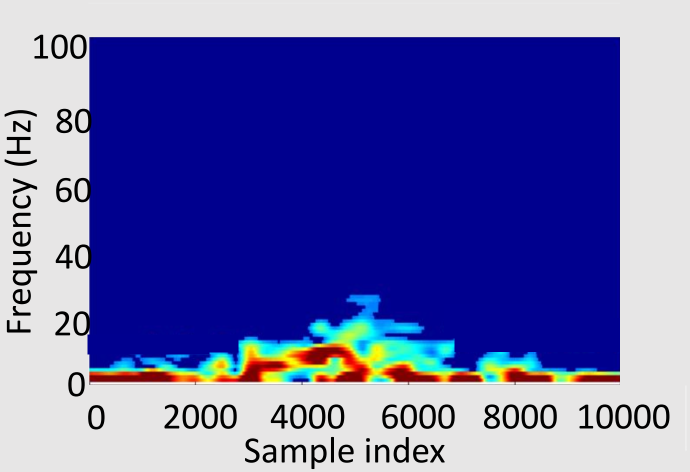


Switching the coding configurations in the desired switch speed can inject controllable interference

Can any two coding configurations achieve effective perturbation?

RISiren - To address Q1

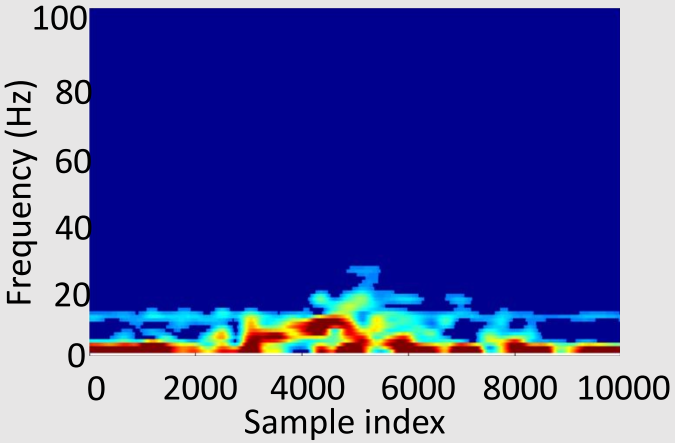
(Q1) How to attack invisibly without extra sources?



Without attack

RISiren - To address Q1

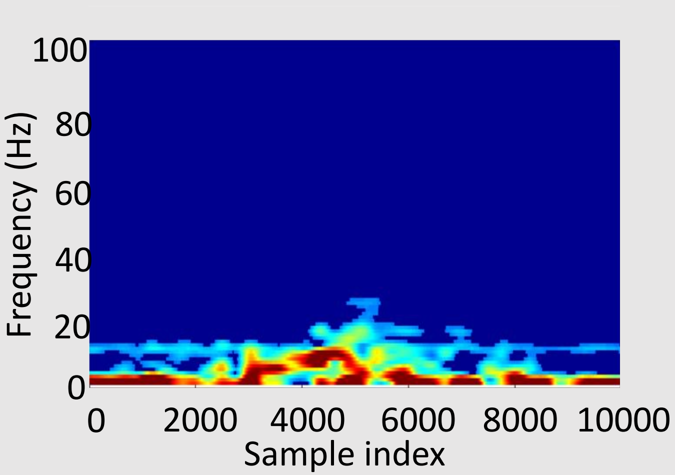
(Q1) How to attack invisibly without extra sources?



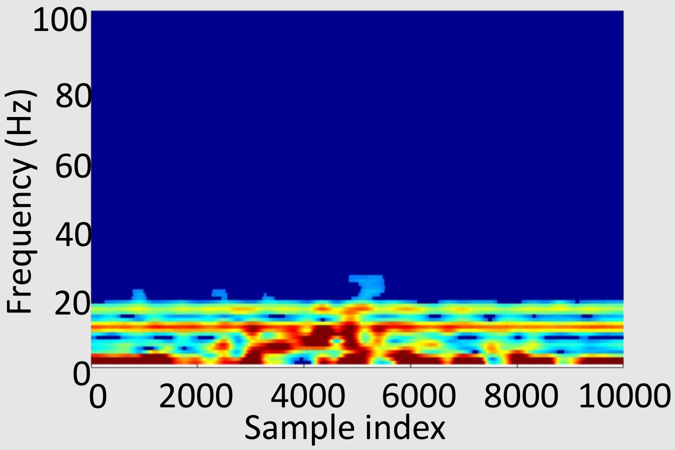
Result in low-intensity attack

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?



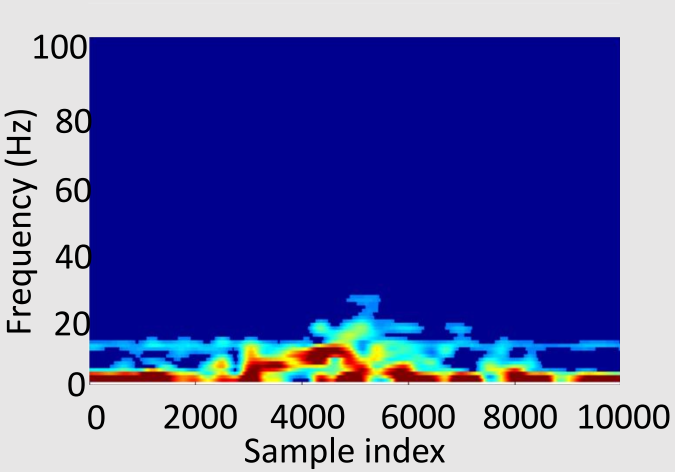
Result in low-intensity attack



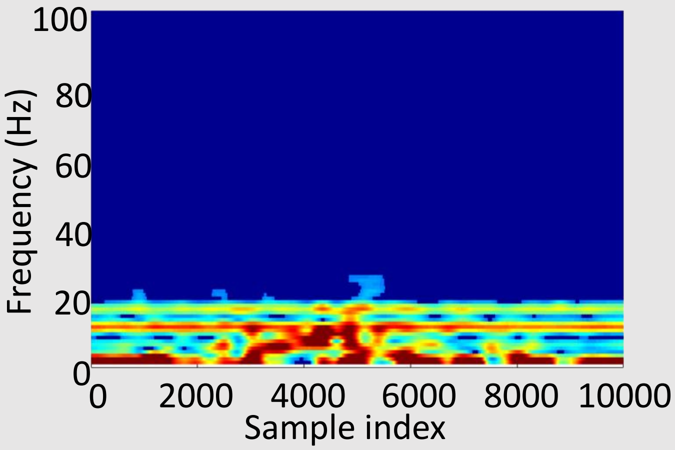
Result in high-intensity attack

RISiren - To address Q1

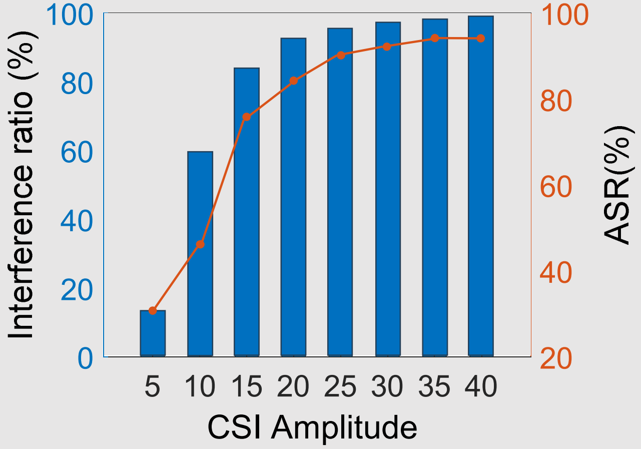
(Q1) How to attack invisibly without extra sources?



Result in low-intensity attack



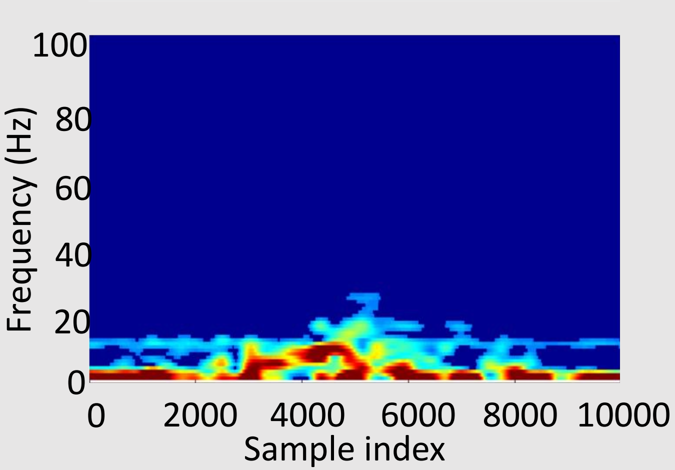
Result in high-intensity attack



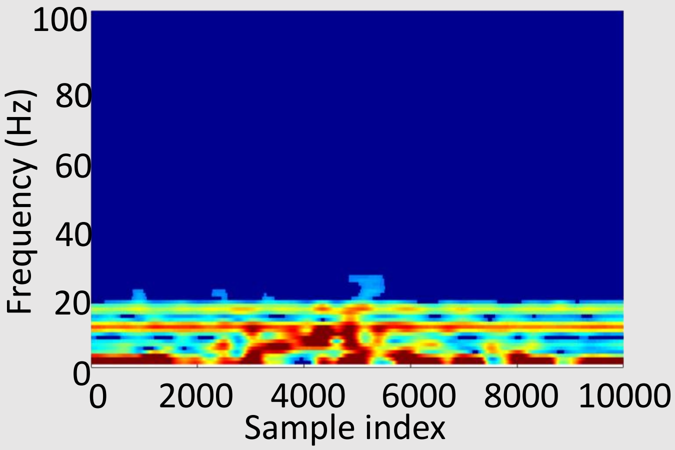
Simulation in different intensity

RISiren - To address Q1

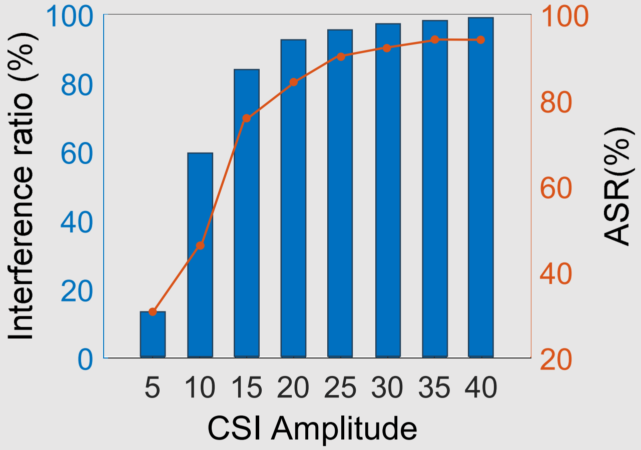
(Q1) How to attack invisibly without extra sources?



Result in low-intensity attack



Result in high-intensity attack



Simulation in different intensity

RISiren should create a high-intensity attack signal to guarantee the effective attacks

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

A straightforward solution:

Coding Configuration1: Beamforming

Coding Configuration2: Metasurface “OFF”

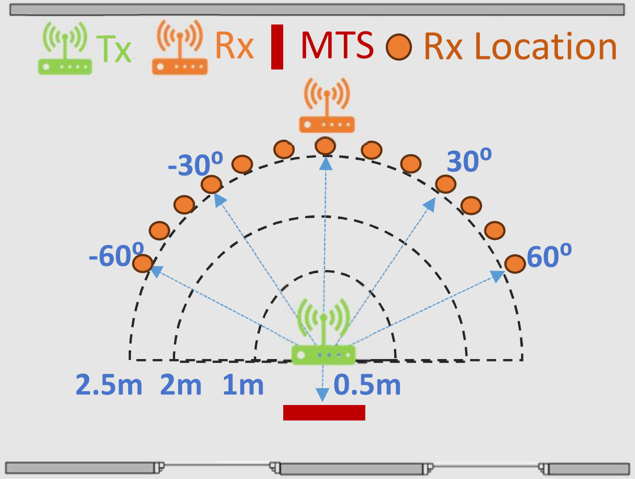
RISiren - To address Q1

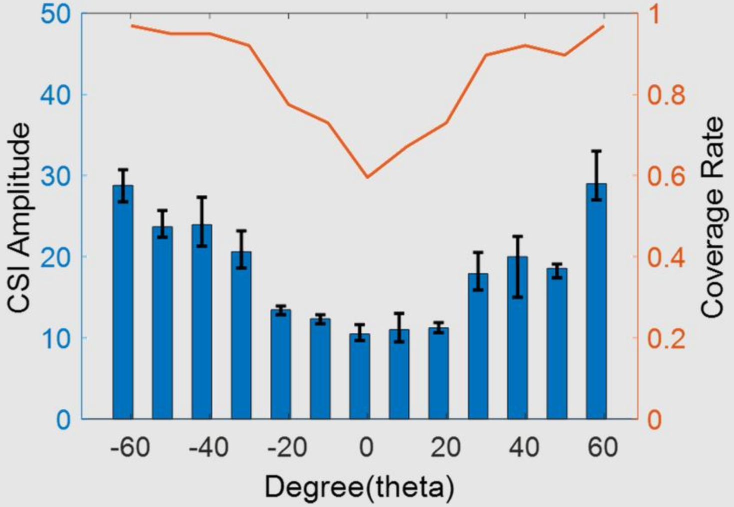
(Q1) How to attack invisibly without extra sources?

A straightforward solution:

Coding Configuration1: Beamforming

Coding Configuration2: Metasurface “OFF”





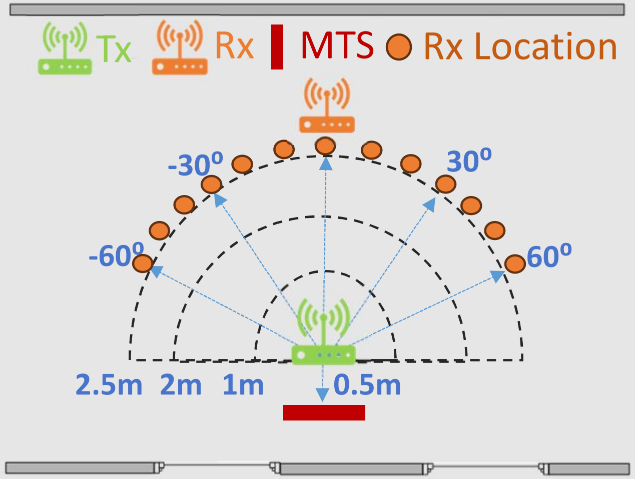
RISiren - To address Q1

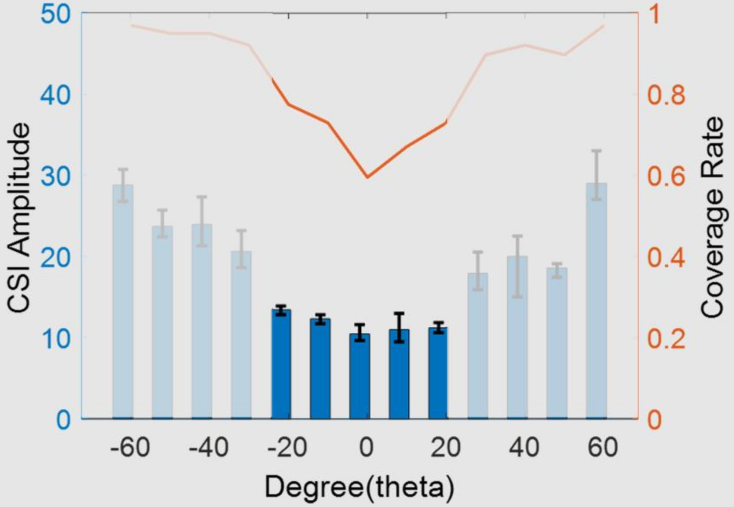
(Q1) How to attack invisibly without extra sources?

A straightforward solution:

Coding Configuration1: Beamforming

Coding Configuration2: Metasurface "OFF"





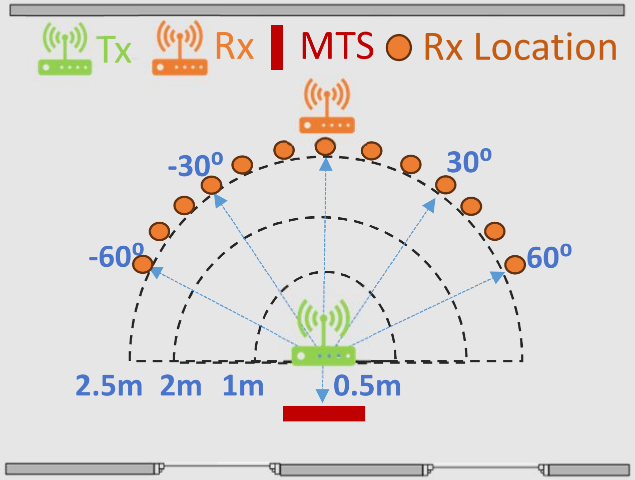
RISiren - To address Q1

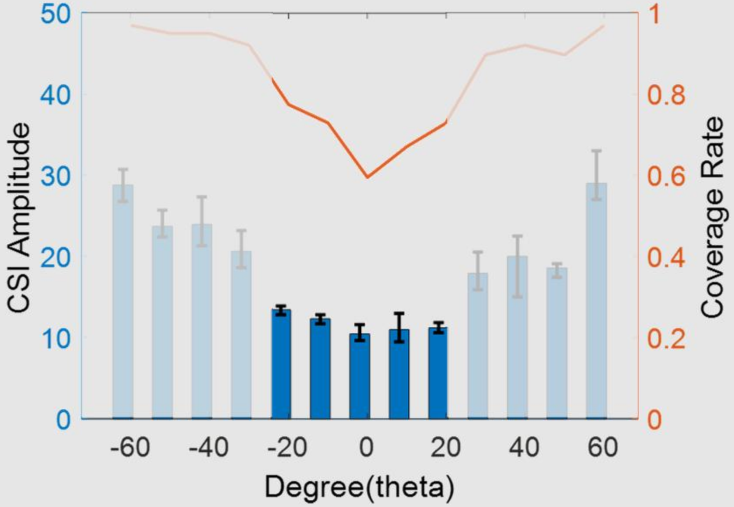
(Q1) How to attack invisibly without extra sources?

A straightforward solution:

Coding Configuration1: Beamforming

Coding Configuration2: Metasurface “OFF”





- ➢ **Due to the signal being reflected by the mirror when the metasurface is turned off near 0°, there is only a minor difference in reflected signal intensity between the two states.**

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

- Optimization algorithm to maximize interference signals -

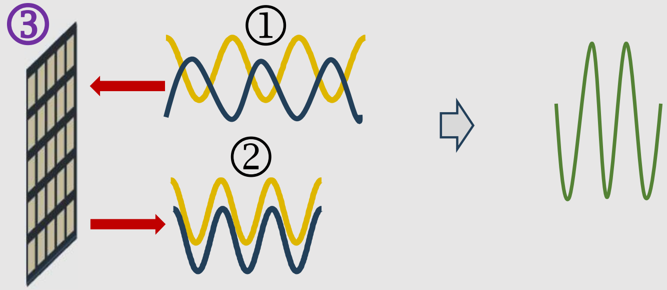
RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

- Optimization algorithm to maximize interference signals -

RISiren solution:

Coding Configuration1: Beamforming



① Incident phase: $\emptyset_{m,n}^{I} = -k_0 d_{m,n}$

② Theoretical phase:

$\emptyset_{m,n}^{T} = -k_0(x_m sin\theta_0 cos\varphi_0 + y_n sin\theta_0 sin\varphi_0)$

③ Compensation phase:

$$\emptyset_{m,n}^{C} = \emptyset_{m,n}^{T} - \emptyset_{m,n}^{I}$$

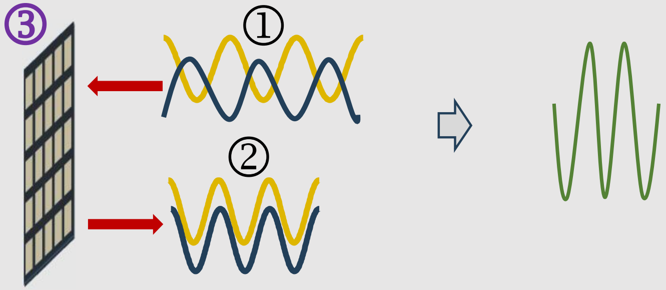
RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

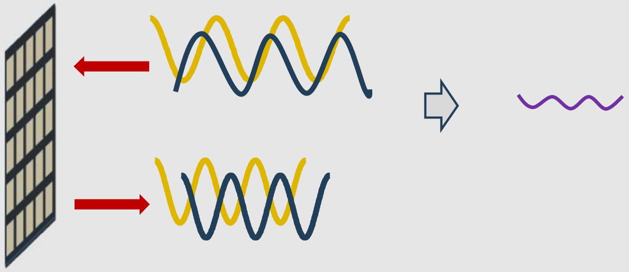
- Optimization algorithm to maximize interference signals -

RISiren solution:

Coding Configuration1: Beamforming



Coding Configuration2: Nullforming



① Incident phase: $\emptyset^{I}_{m,n} = -k_0 d_{m,n}$

② Theoretical phase:

$$\emptyset^{T}_{m,n} = -k_0(x_m sin\theta_0 cos\varphi_0 + y_n sin\theta_0 sin\varphi_0)$$

③ **Compensation phase:**

$$\emptyset^{C}_{m,n} = \emptyset^{T}_{m,n} - \emptyset^{I}_{m,n}$$

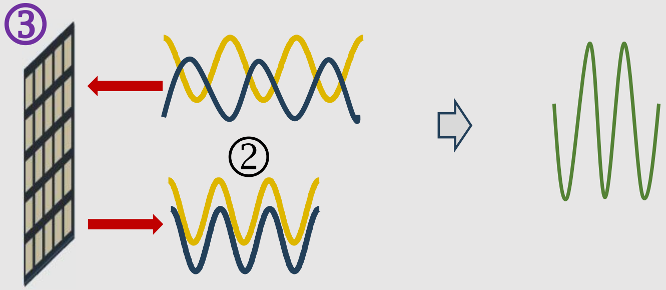
RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

- Optimization algorithm to maximize interference signals -

RISiren solution:

Coding Configuration1: Beamforming



① Incident phase: $\emptyset_{m,n}^{I} = -k_0 d_{m,n}$

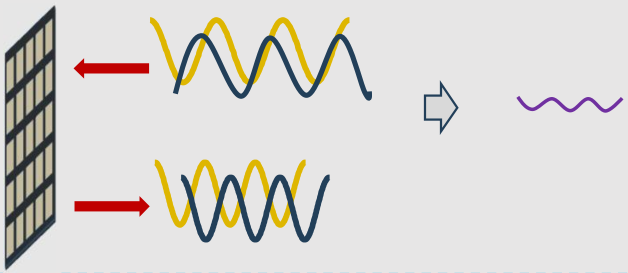
② Theoretical phase:

$$\emptyset_{m,n}^{T} = -k_0(x_m sin\theta_0 cos\varphi_0 + y_n sin\theta_0 sin\varphi_0)$$

③ Compensation phase:

$$\emptyset_{m,n}^{C} = \emptyset_{m,n}^{T} - \emptyset_{m,n}^{I}$$

Coding Configuration2: Nullforming



$$\mathcal{L} \in \min \sqrt{\ell_1^2 + \ell_2^2 + \ell_3^2}$$

Ensure the Nullfrorming gain: $\ell_1 = |Gain_{(\theta_\ell,\varphi_\ell)} - BFGain_{(\theta_\ell,\varphi_\ell)}|^{-1}$

$S.t$

$$(\theta_\ell,\varphi_\ell) \in [(\theta_\ell,\varphi_\ell) - \frac{BW_1}{2}, (\theta_\ell,\varphi_\ell) + \frac{BW_1}{2}]$$

$$\gamma \in C_u^{((\theta_\ell,\varphi_\ell))}$$

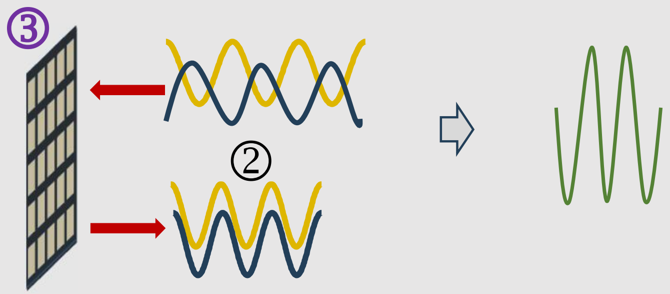
RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

- Optimization algorithm to maximize interference signals -

RISiren solution:

Coding Configuration1: Beamforming



① Incident phase: $\emptyset^{I}_{m,n} = -k_0 d_{m,n}$

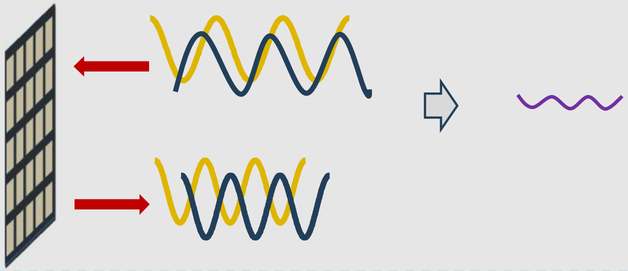
② Theoretical phase:

$$\emptyset^{T}_{m,n} = -k_0(x_m sin\theta_0 cos\varphi_0 + y_n sin\theta_0 sin\varphi_0)$$

③ **Compensation phase:**

$$\emptyset^{C}_{m,n} = \emptyset^{T}_{m,n} - \emptyset^{I}_{m,n}$$

Coding Configuration2: Nullforming



$$\mathcal{L} \in \min \sqrt{\ell_1^2 + \ell_2^2 + \ell_3^2}$$

Ensure the Nullfrorming gain: $\ell_1 = |Gain_{(\theta_\ell,\varphi_\ell)} - BFGain_{(\theta_\ell,\varphi_\ell)}|^{-1}$

Ensure the beam flatness: $\ell_2 = Var(Gain_{(\theta_\ell,\varphi_\ell)})$

$S.t$

$$(\theta_\ell,\varphi_\ell) \in [(\theta_\ell,\varphi_\ell) - \frac{BW_1}{2}, (\theta_\ell,\varphi_\ell) + \frac{BW_1}{2}]$$

$$\gamma \in C_u^{((\theta_\ell,\varphi_\ell))}$$

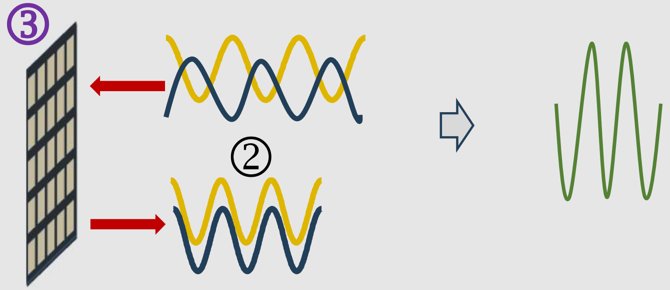
RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

- Optimization algorithm to maximize interference signals -

RISiren solution:

Coding Configuration1: Beamforming



① Incident phase: $\emptyset^{I}_{m,n} = -k_0 d_{m,n}$

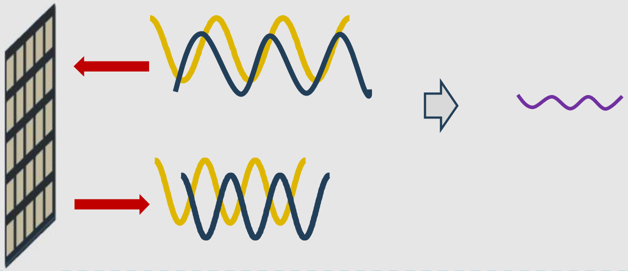
② Theoretical phase:

$$\emptyset^{T}_{m,n} = -k_0(x_m sin\theta_0 cos\varphi_0 + y_n sin\theta_0 sin\varphi_0)$$

③ Compensation phase:

$$\emptyset^{C}_{m,n} = \emptyset^{T}_{m,n} - \emptyset^{I}_{m,n}$$

Coding Configuration2: Nullforming



$$\mathcal{L} \in \min\sqrt{\ell_1^2 + \ell_2^2 + \ell_3^2}$$

Ensure the Nullfrorming gain: $\ell_1 = |Gain_{(\theta_\ell,\varphi_\ell)} - BFGain_{(\theta_\ell,\varphi_\ell)}|^{-1}$

Ensure the beam flatness: $\ell_2 = Var(Gain_{(\theta_\ell,\varphi_\ell)})$

Ensure the sidelobe gain: $\ell_3 = Max(Gain_\gamma) - Min(Gain_\gamma)$

$S.t$

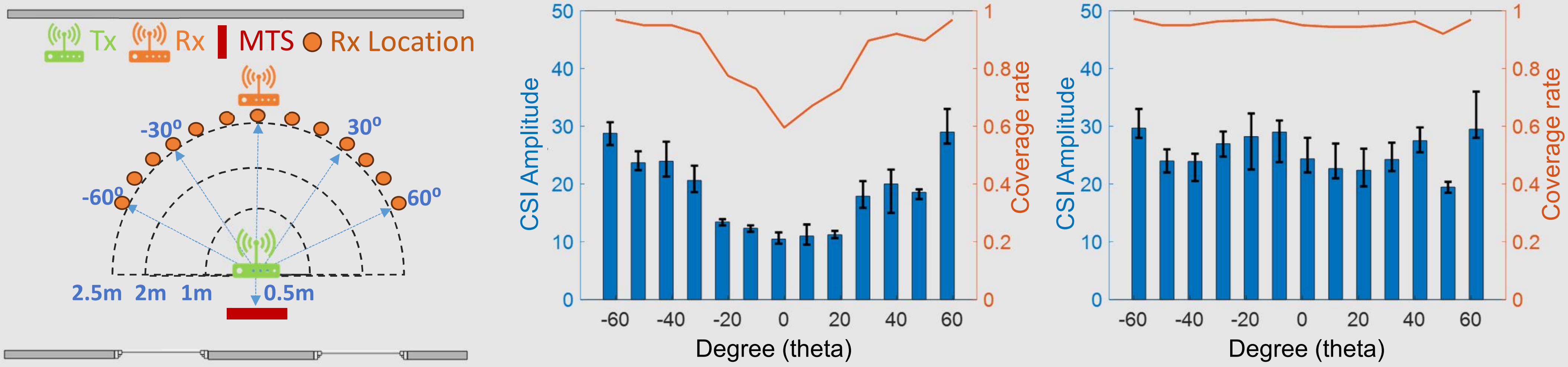
$$(\theta_\ell,\varphi_\ell) \in [(\theta_\ell,\varphi_\ell) - \frac{BW_1}{2}, (\theta_\ell,\varphi_\ell) + \frac{BW_1}{2}]$$

$$\gamma \in C_u^{((\theta_\ell,\varphi_\ell))}$$

RISiren - To address Q1

(Q1) How to attack invisibly without extra sources?

- Optimization algorithm to maximize interference signals -



(a) The experiment layout

(b) Beamforming and metasurface “OFF”

(c) Beamforming and nullforming

Our work RISiren[1]

Q1: How to inject the adversarial attack invisibly?

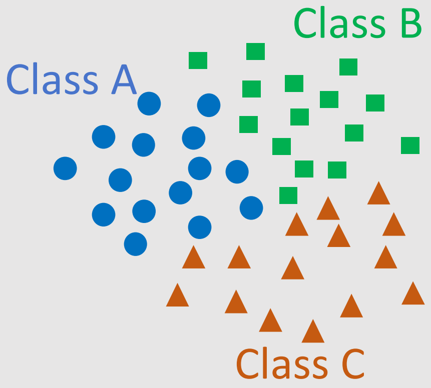
Q2: How to achieve a black-box attack?

[1] “RISiren” derived from the sea-nymphs “Siren” who lured sailors to their death with a bewitching song in ancient Greek mythology

RISiren - To address Q2

(Q2) How to achieve a black-box attack?

Prior solution analyzation in feature domain

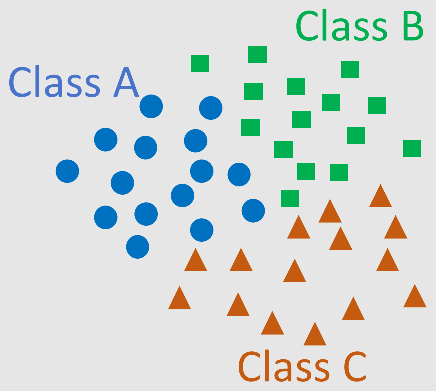


Feature distribution in feature-domain

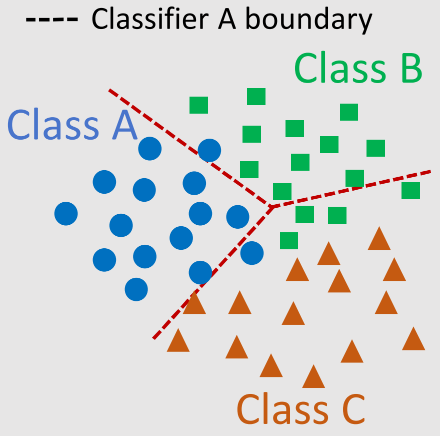
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

Prior solution analyzation in feature domain



Feature distribution in feature-domain

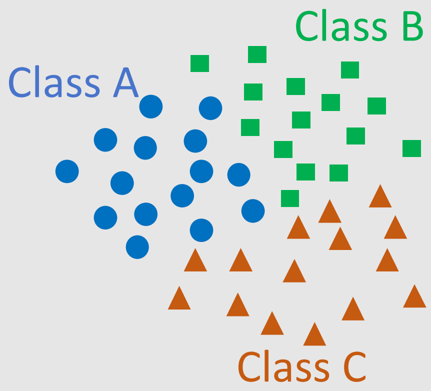


Get the victim classifier architecture

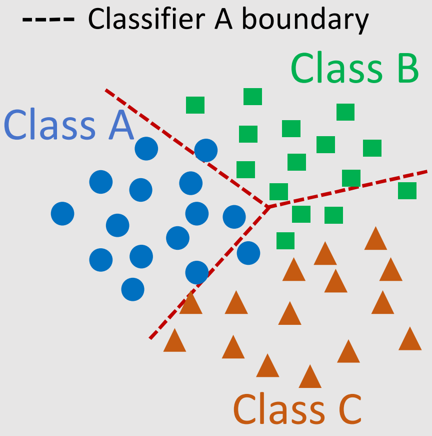
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

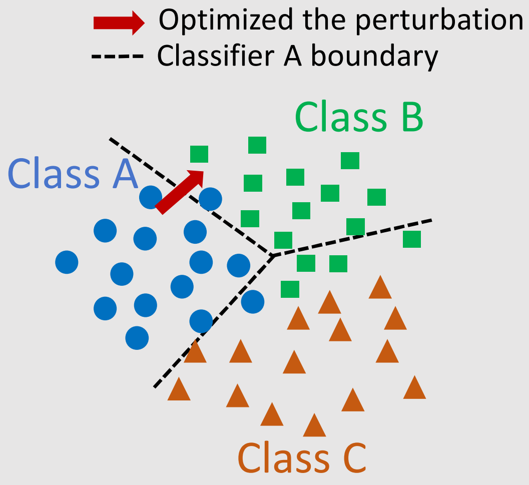
Prior solution analyzation in feature domain



Feature distribution in feature-domain



Get the victim classifier architecture

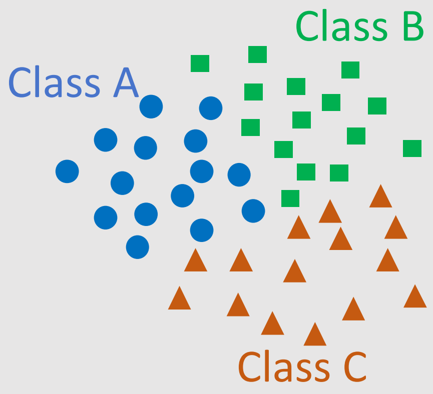


optimize the perturbation to cross the boundary

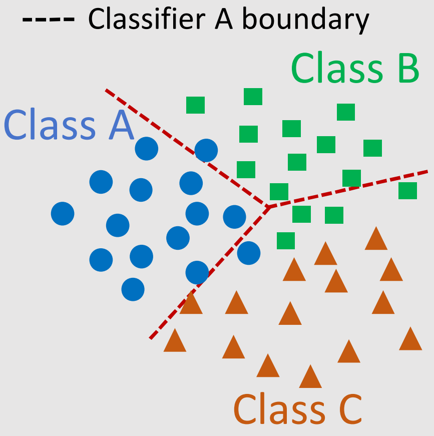
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

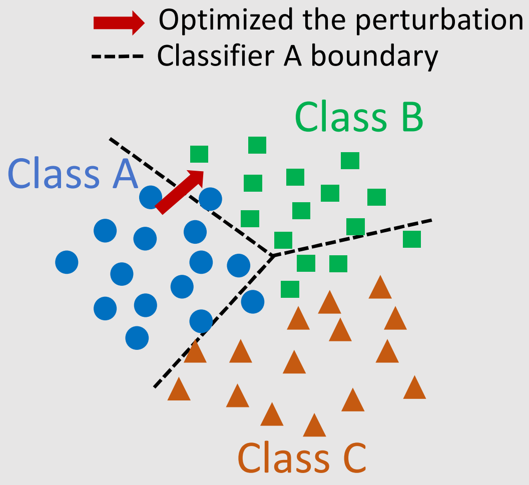
Prior solution analyzation in feature domain



Feature distribution in feature-domain



Get the victim classifier architecture



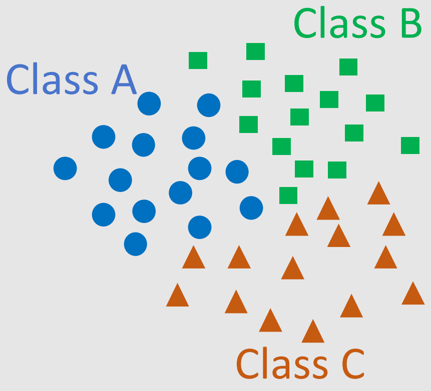
optimize the perturbation to cross the boundary

Limitation1: The victim classifier architecture is hard to get in the physical attack

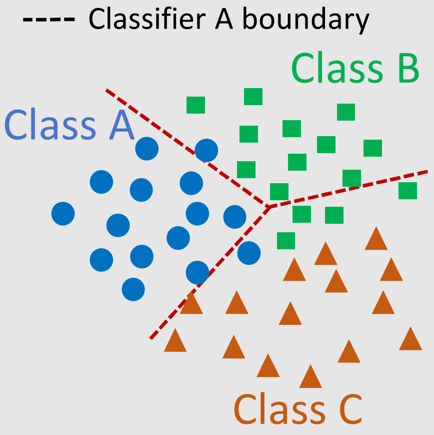
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

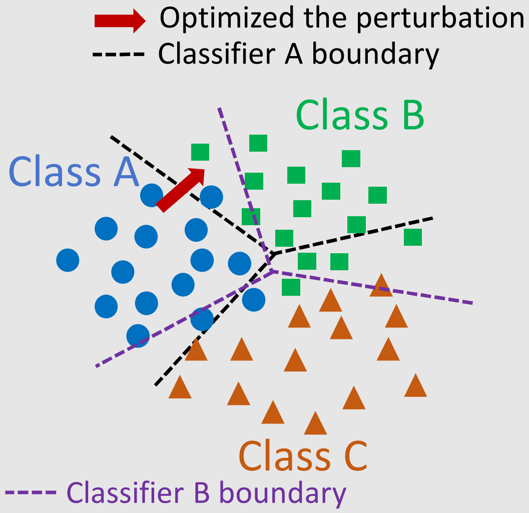
Prior solution analyzation in feature domain



Feature distribution in feature-domain



Get the victim classifier architecture


Optimized the perturbation
---- Classifier A boundary
Class B
Class A
Class C
---- Classifier B boundary

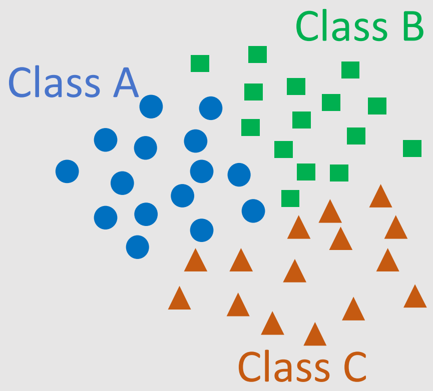
optimize the perturbation to cross the boundary

Limitation1: The victim classifier architecture is hard to get in the physical attack

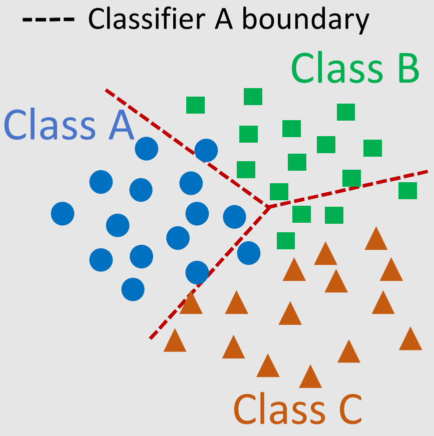
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

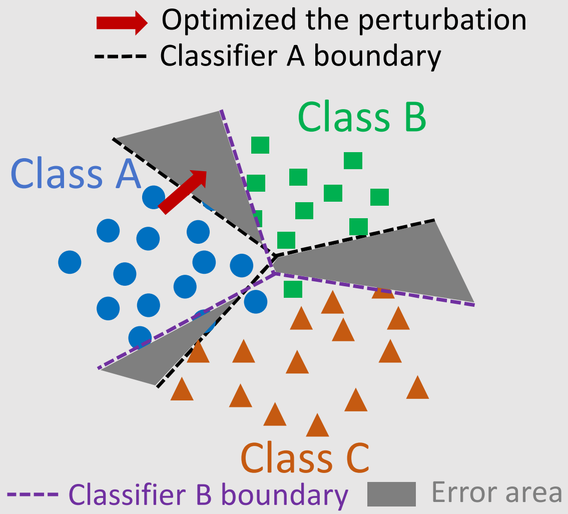
Prior solution analyzation in feature domain



Feature distribution in feature-domain



Get the victim classifier architecture


Optimized the perturbation
---- Classifier A boundary
Class A
Class B
Class C
---- Classifier B boundary
Error area

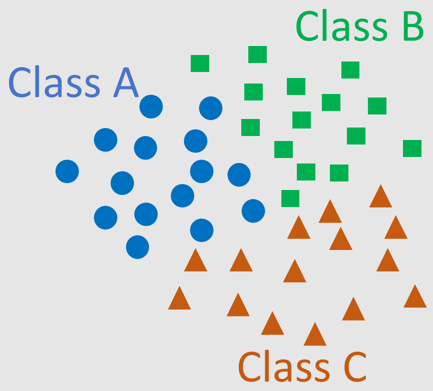
optimize the perturbation to cross the boundary

Limitation1: The victim classifier architecture is hard to get in the physical attack

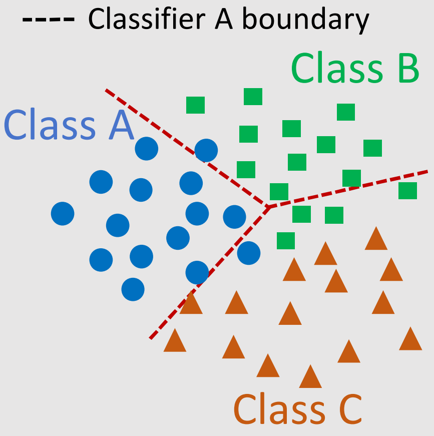
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

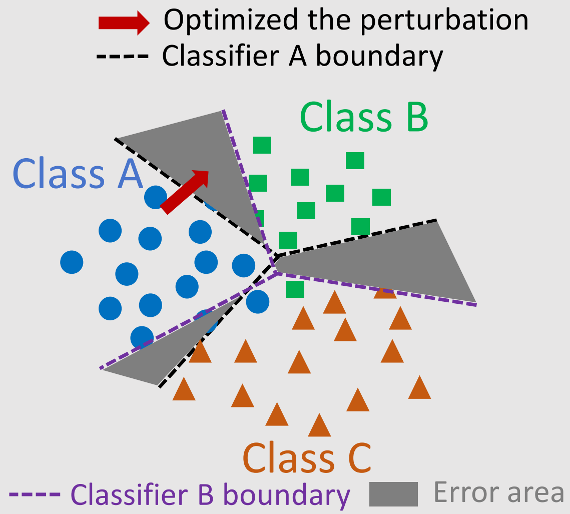
Prior solution analyzation in feature domain



Feature distribution in feature-domain



Get the victim classifier architecture


Optimized the perturbation
---- Classifier A boundary
Class B
Class A
Class C
---- Classifier B boundary
Error area

optimize the perturbation to cross the boundary

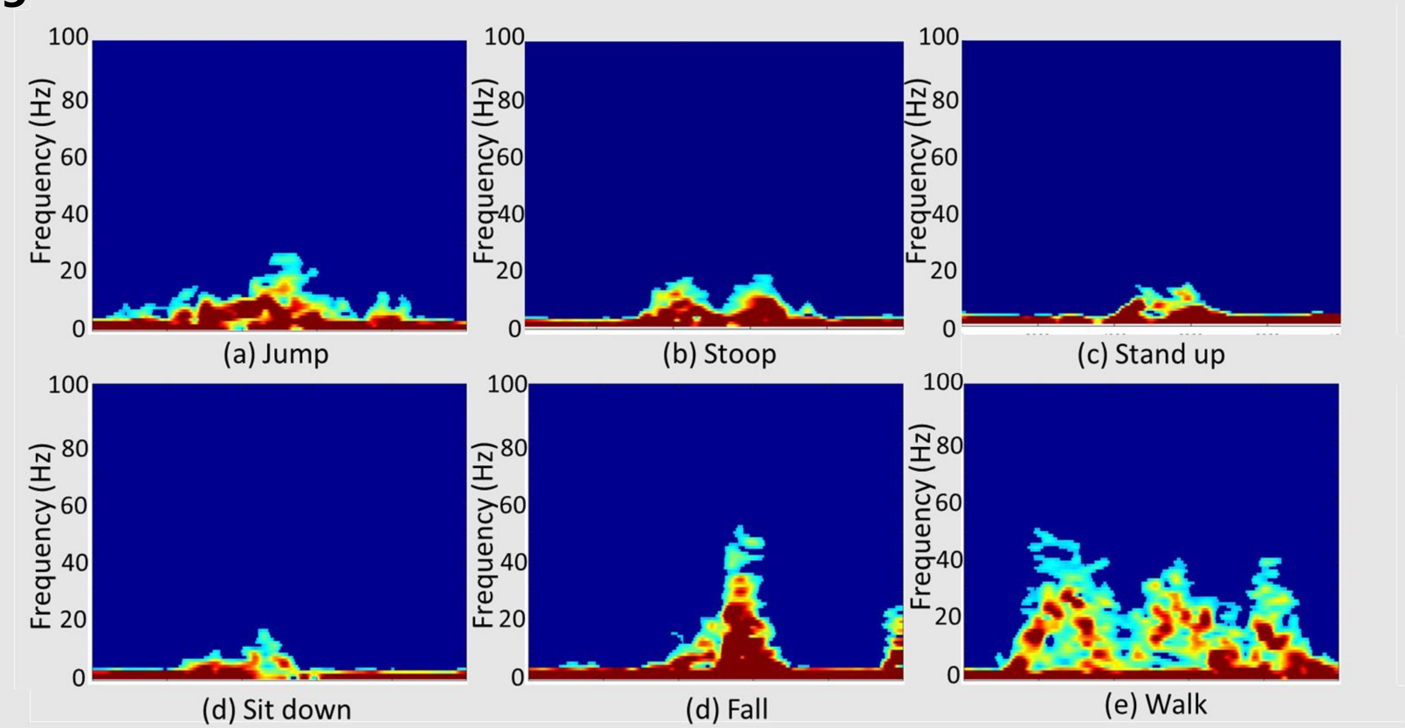
Limitation1: The victim classifier architecture is hard to get in the physical attack

Limitation2: Adversarial perturbations have low generalization performance

RISiren - To address Q2

(Q2) How to achieve a black-box attack?

An interesting observation:



(a) Jump

(b) Stoop

(c) Stand up

(d) Sit down

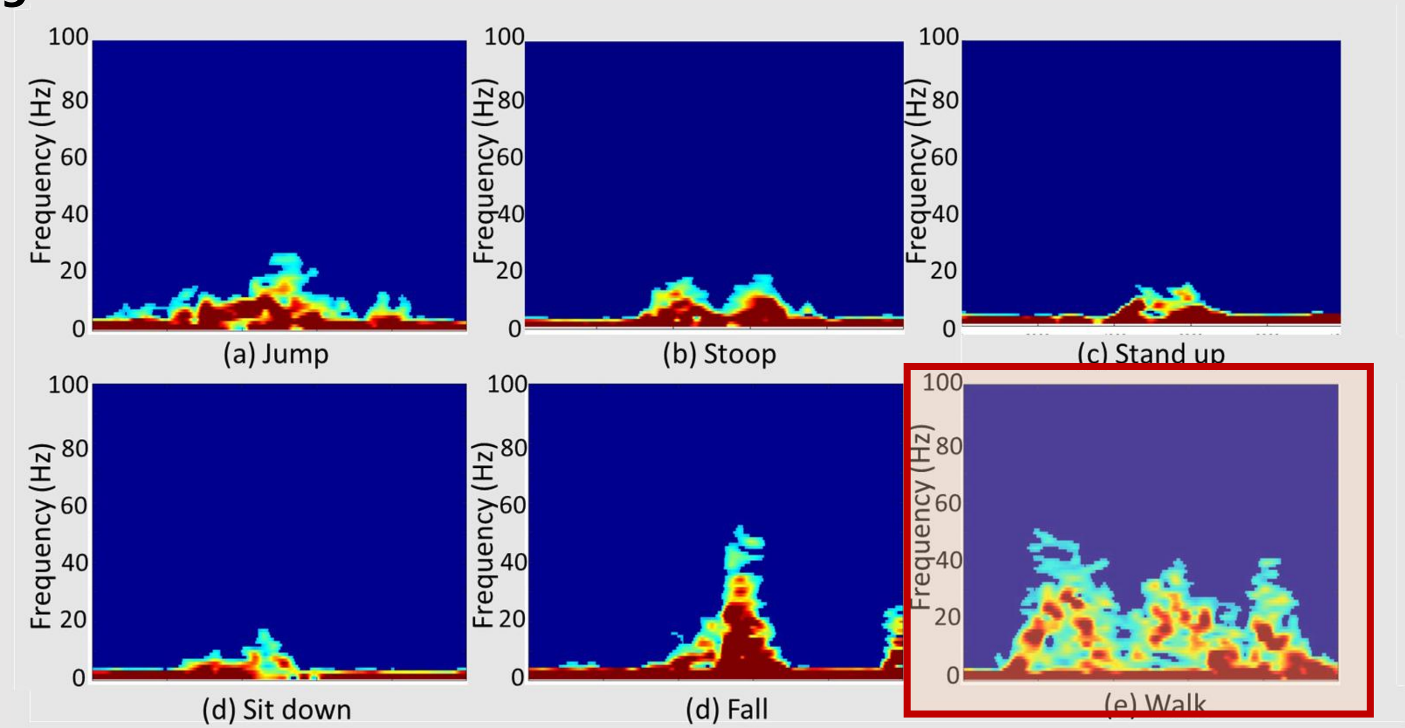
(d) Fall

(e) Walk

RISiren - To address Q2

(Q2) How to achieve a black-box attack?

An interesting observation:



(a) Jump

(b) Stoop

(c) Stand up

(d) Sit down

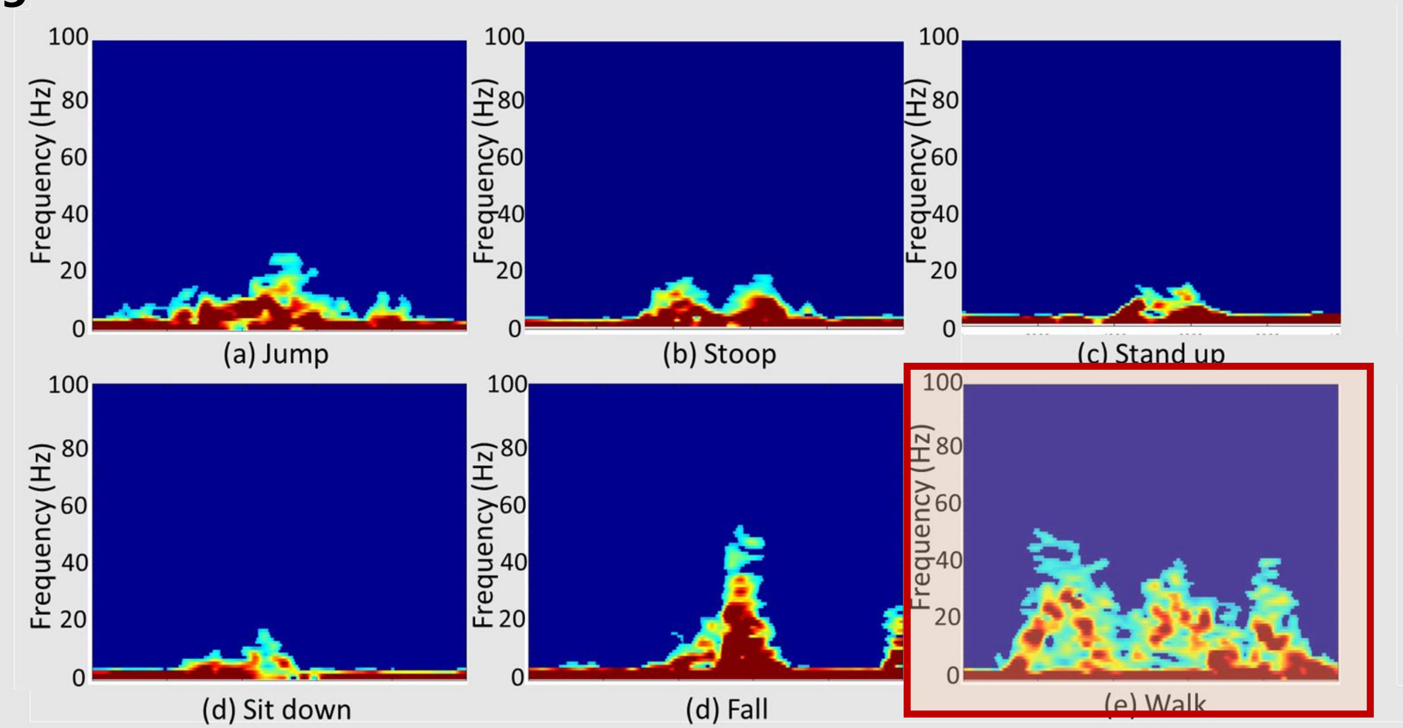
(d) Fall

(e) Walk

RISiren - To address Q2

(Q2) How to achieve a black-box attack?

An interesting observation:



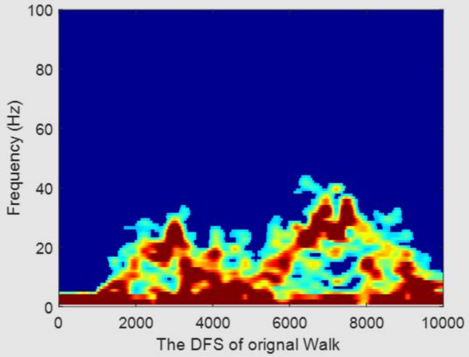
We can generate a carefully designed robust fake activity feature to mask the original activity feature

RISiren - To address Q2

(Q2) How to achieve a black-box attack?

- A camouflaged activity framework -

Step1: Extrat the frequency spectrum outline cure

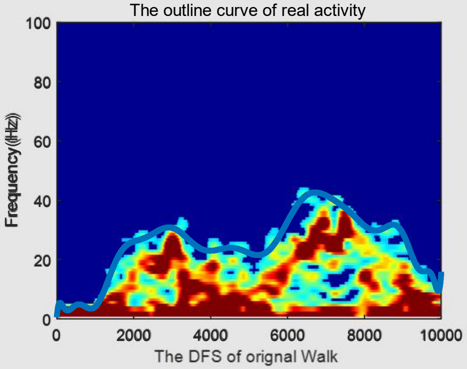


RISiren - To address Q2

(Q2) How to achieve a black-box attack?

- A camouflaged activity framework -

Step1: Extrat the frequency spectrum outline cure

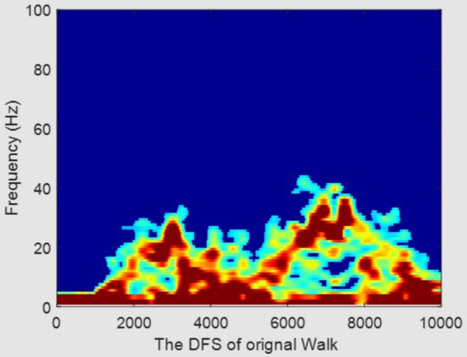


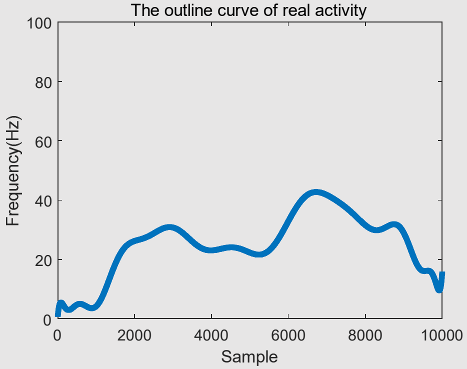
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

- A camouflaged activity framework -

Step1: Extrat the frequency spectrum outline cure



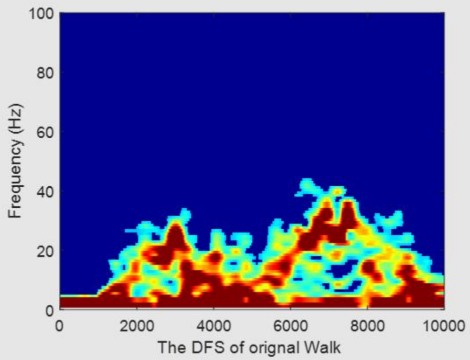


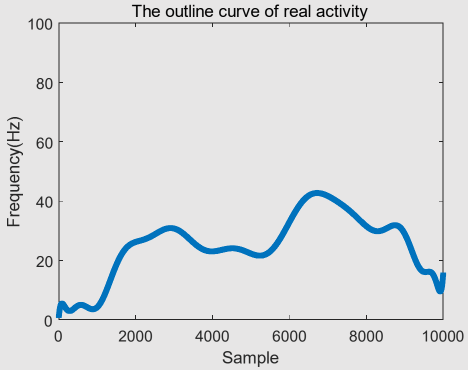
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

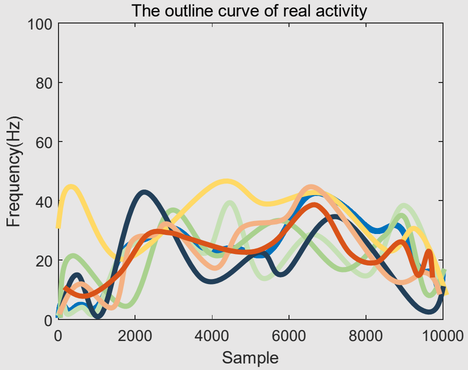
- A camouflaged activity framework -

Step1: Extrat the frequency spectrum outline cure





Step2: an approximation algorithm to fit the truth cure

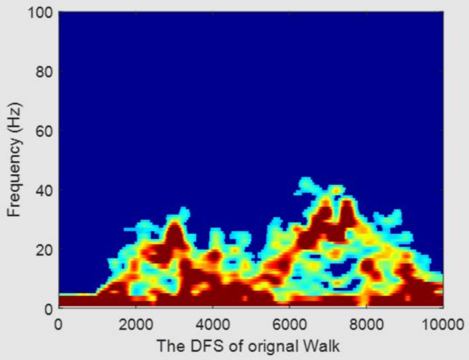


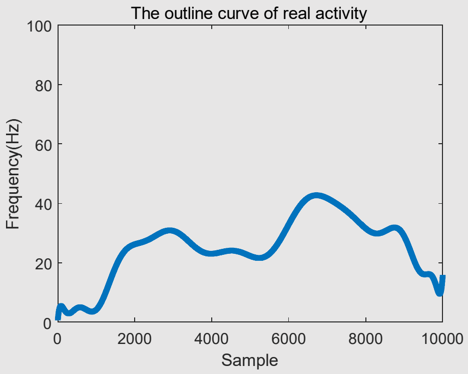
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

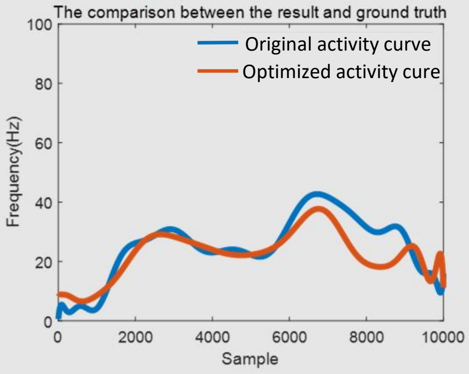
- A camouflaged activity framework -

Step1: Extrat the frequency spectrum outline cure





Step2: an approximation algorithm to fit the truth cure

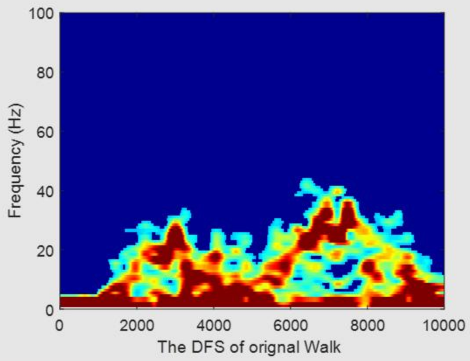


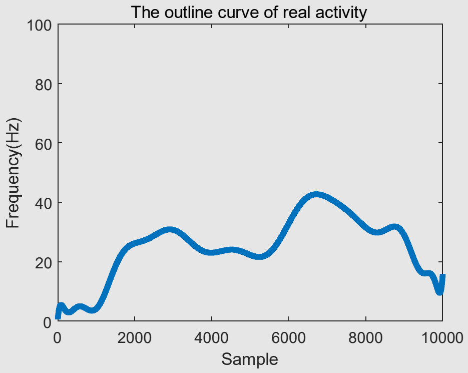
RISiren - To address Q2

(Q2) How to achieve a black-box attack?

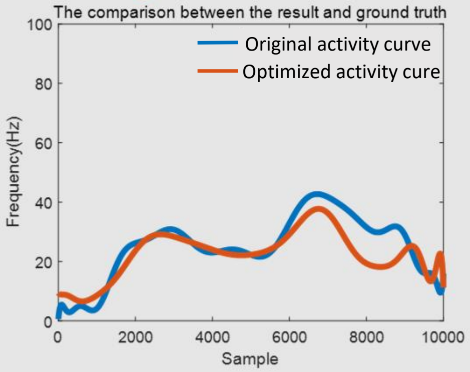
- A camouflaged activity framework -

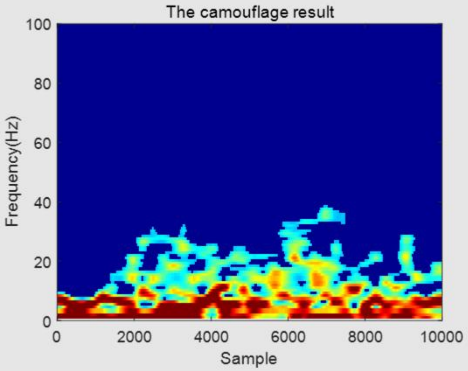
Step1: Extrat the frequency spectrum outline cure



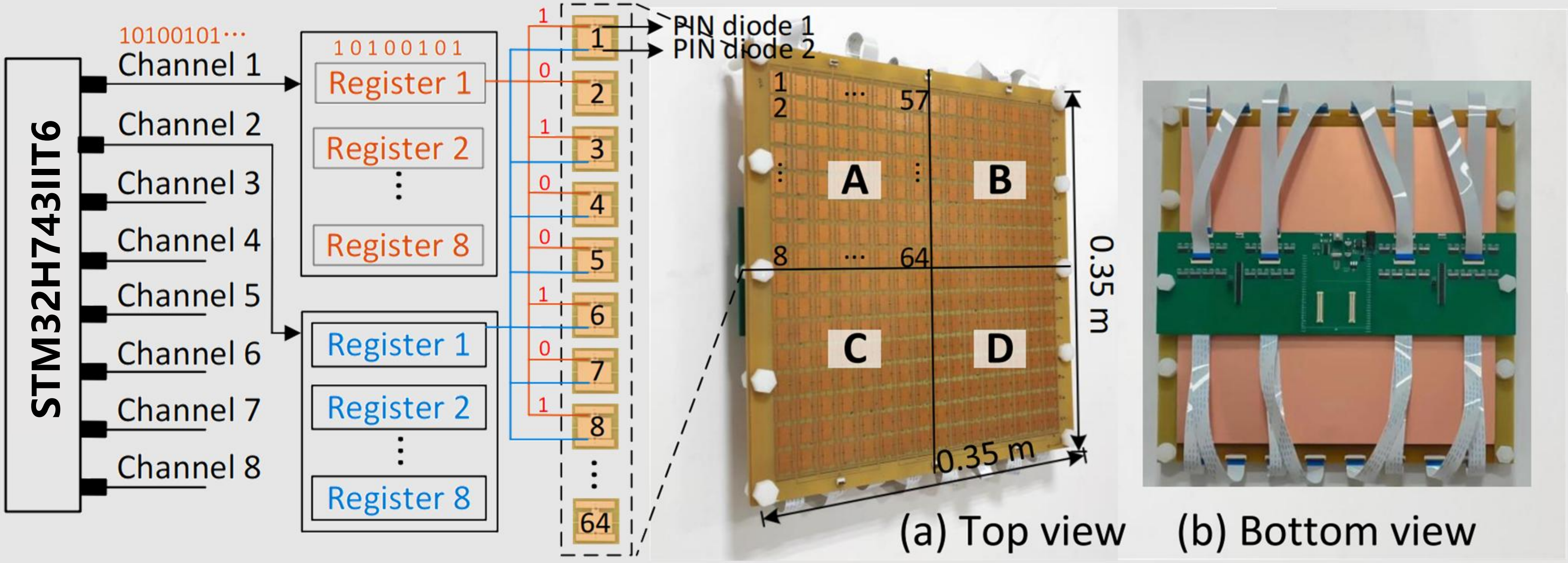


Step2: an approximation algorithm to fit the truth cure





Implementation



(a) Top view (b) Bottom view

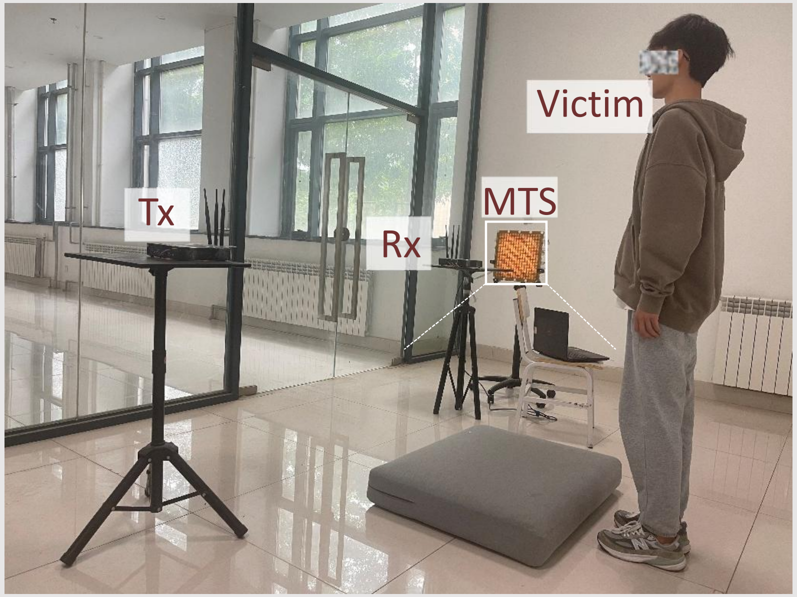
Size: including 256 meta-atoms, area is 35×35 cm^2,thickness is 6.8mm

Control: STM32H743IIT6 controllers and 64 SN74LV595 shift registers.

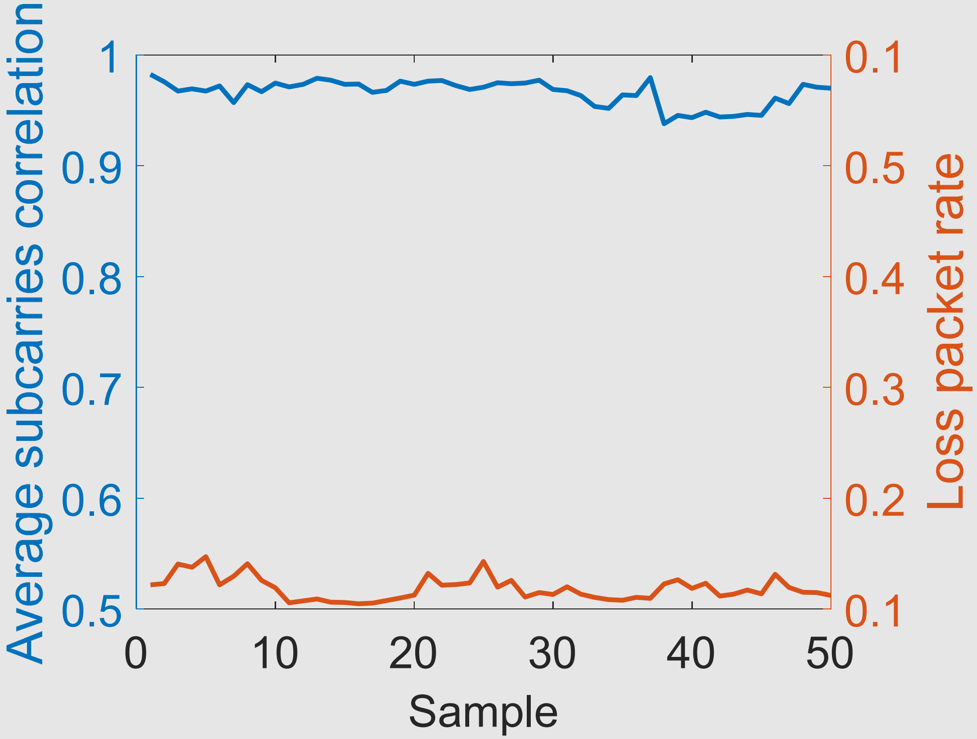
Frequency Support: 2.4GHz & 5GHz

Evaluation

Performance of stealthiness



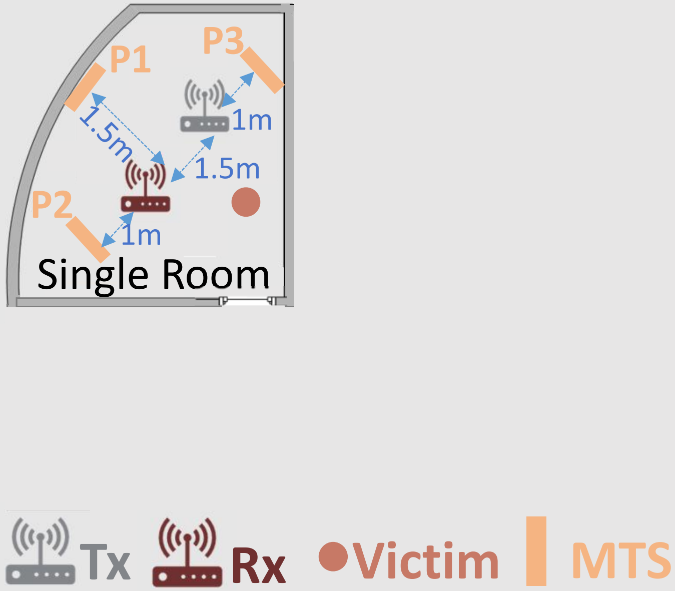
(a) Experiment scenario



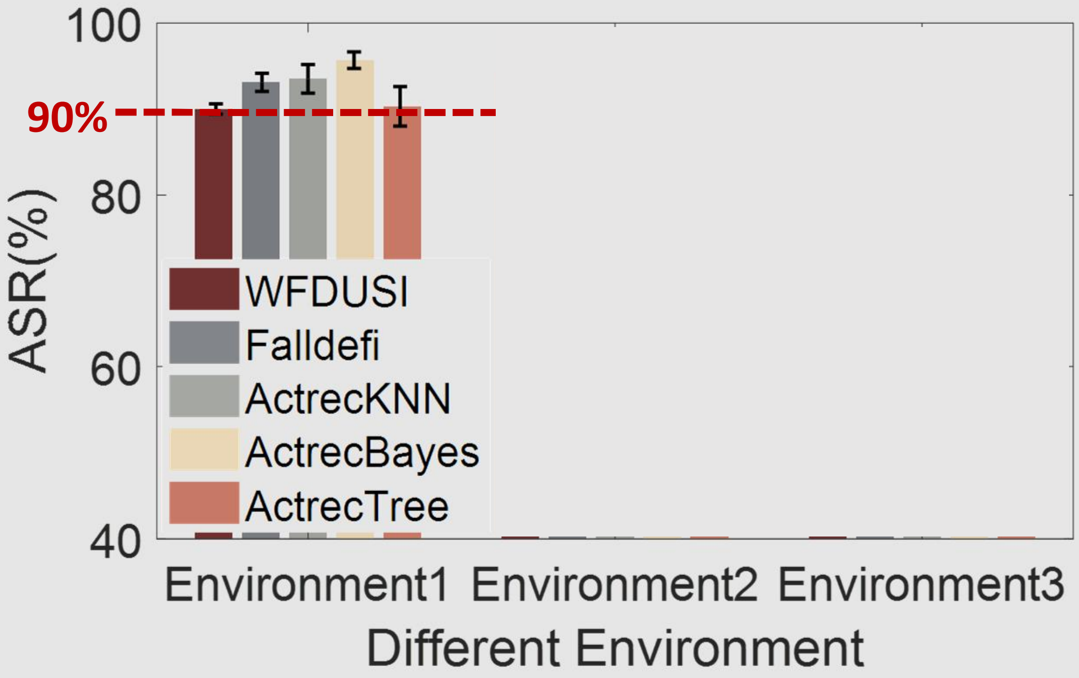
RISiren remains stealthy and hard to detect during attack.

Evaluation

Performance under different environments

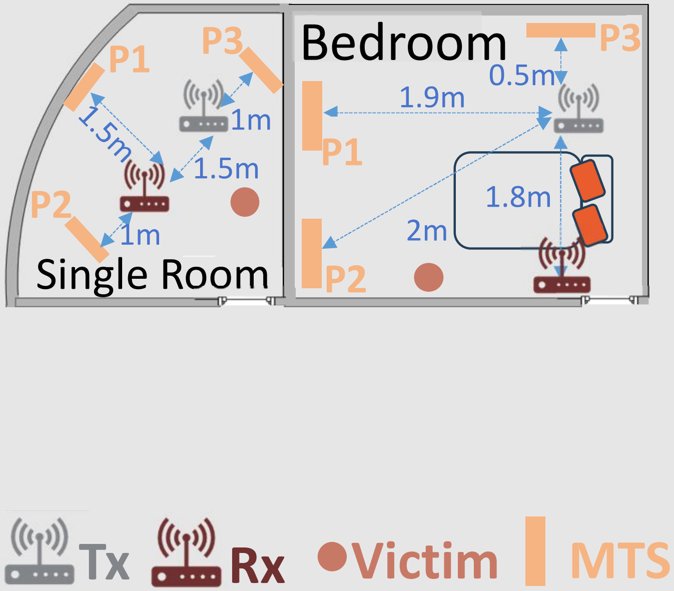


(a) The Scenario layout

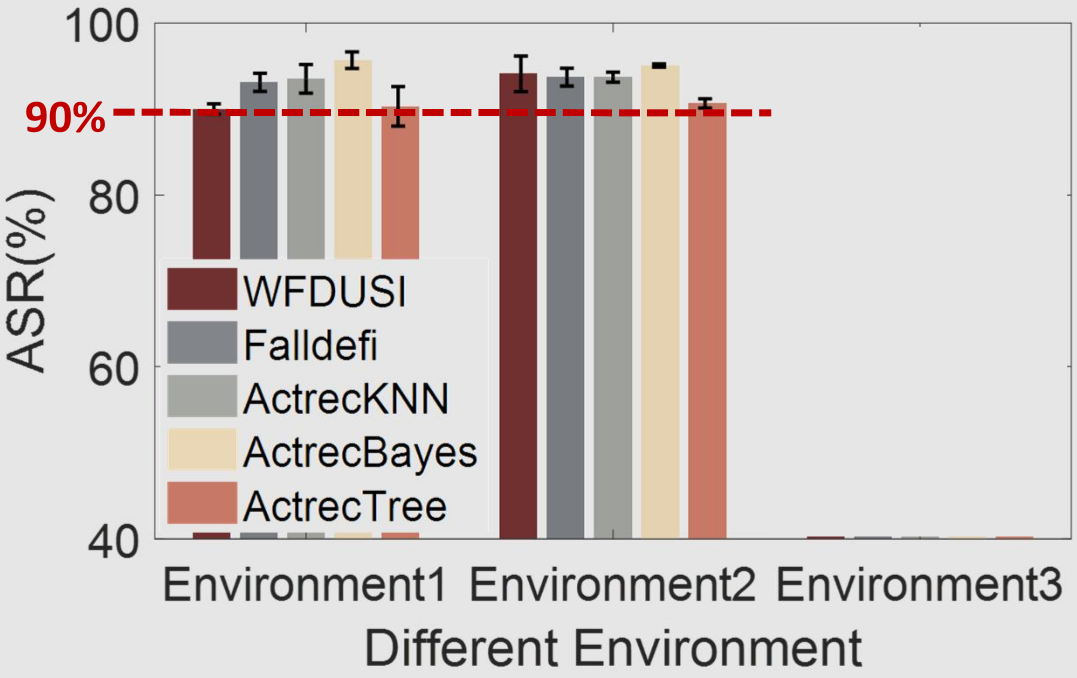


Evaluation

Performance under different environments

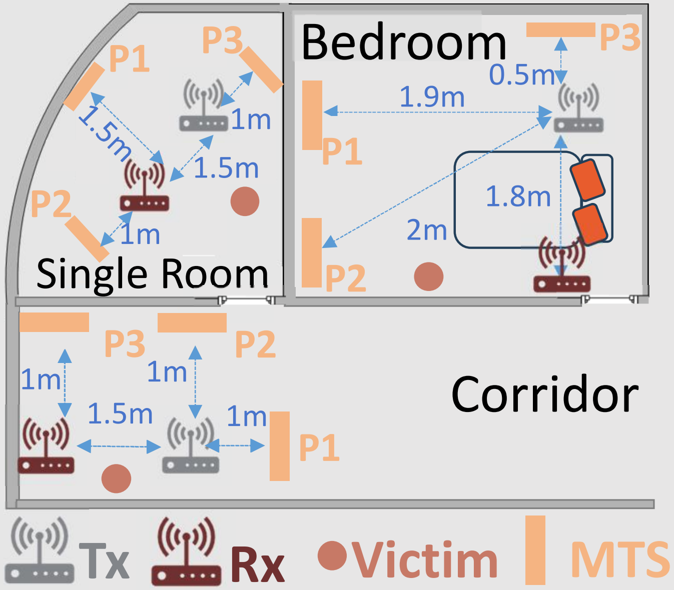


(a) The Scenario layout

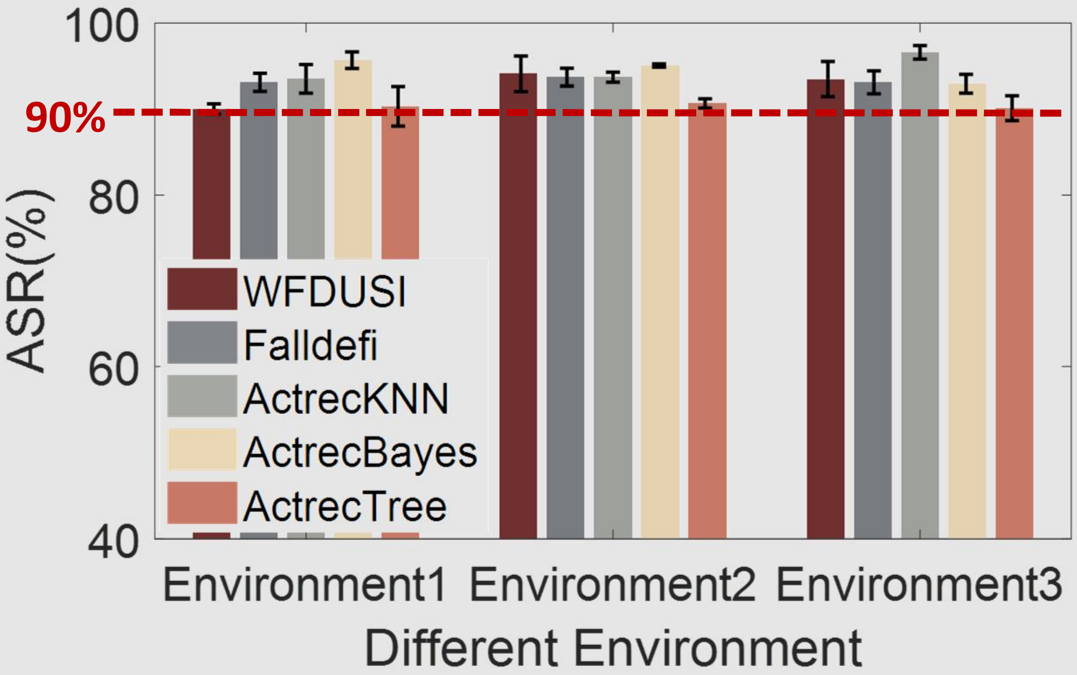


Evaluation

Performance under different environments

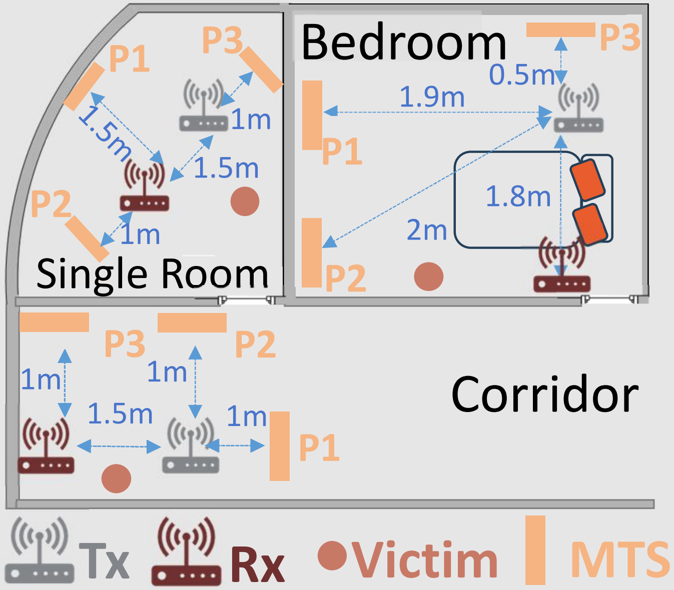


(a) The Scenario layout

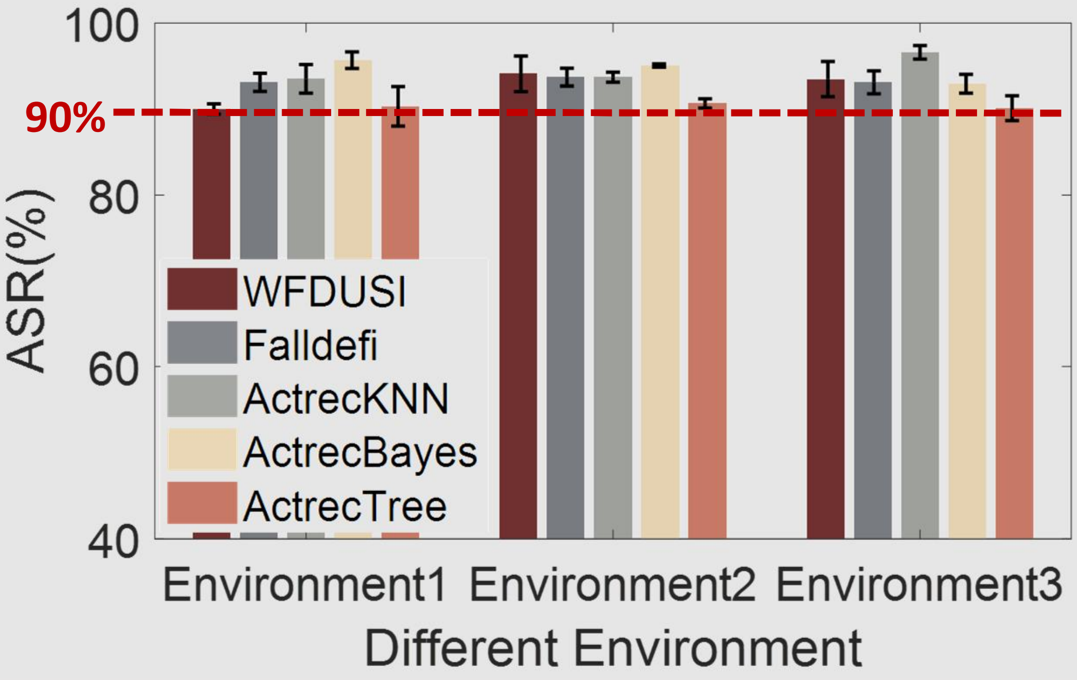


Evaluation

Performance under different environments



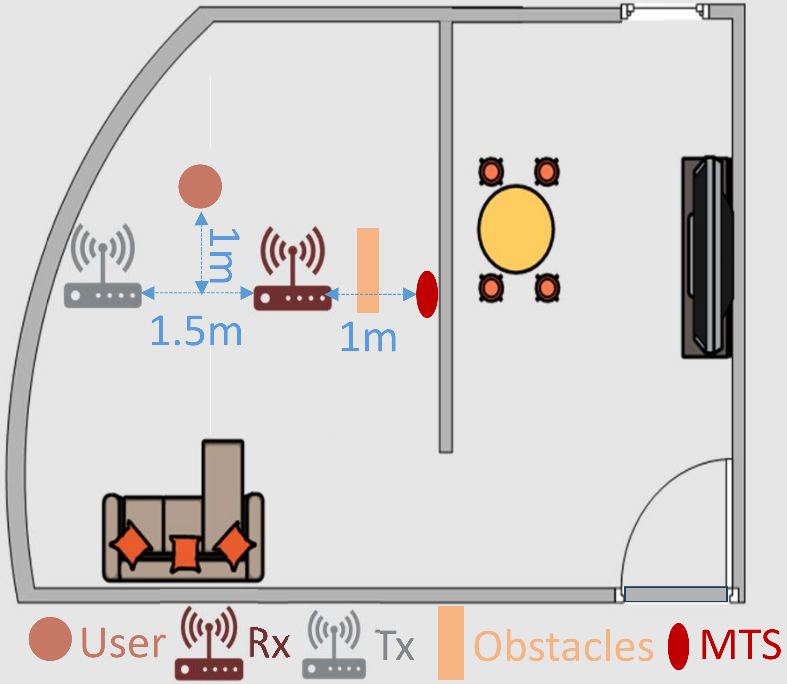
(a) The Scenario layout

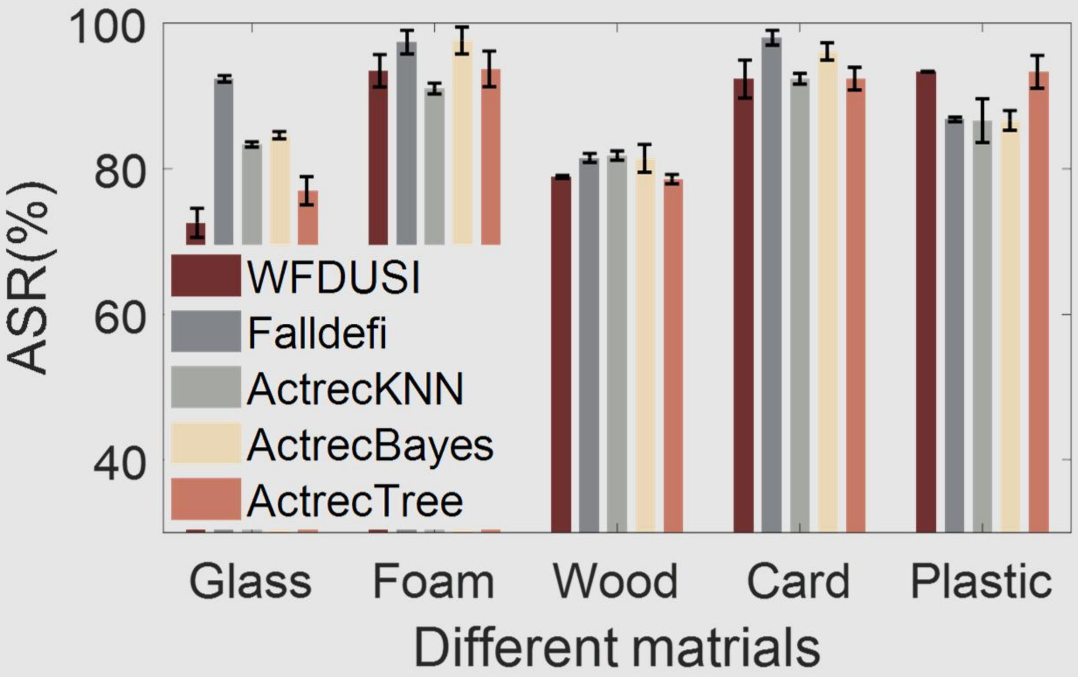


RISiren is robust to the environment

Evaluation

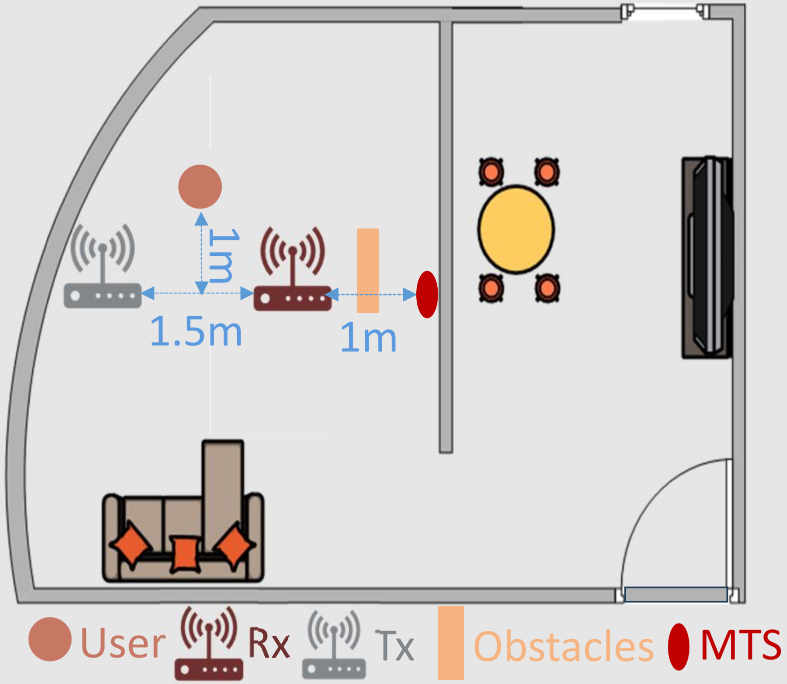
Performance under different obstacles.

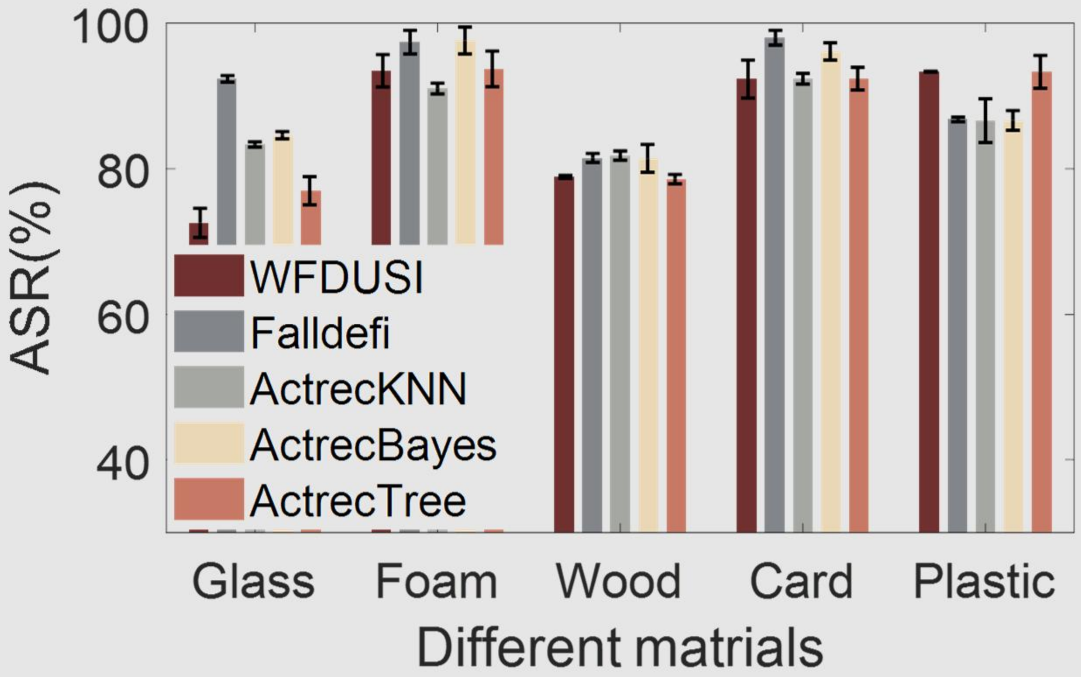




Evaluation

Performance under different obstacles.

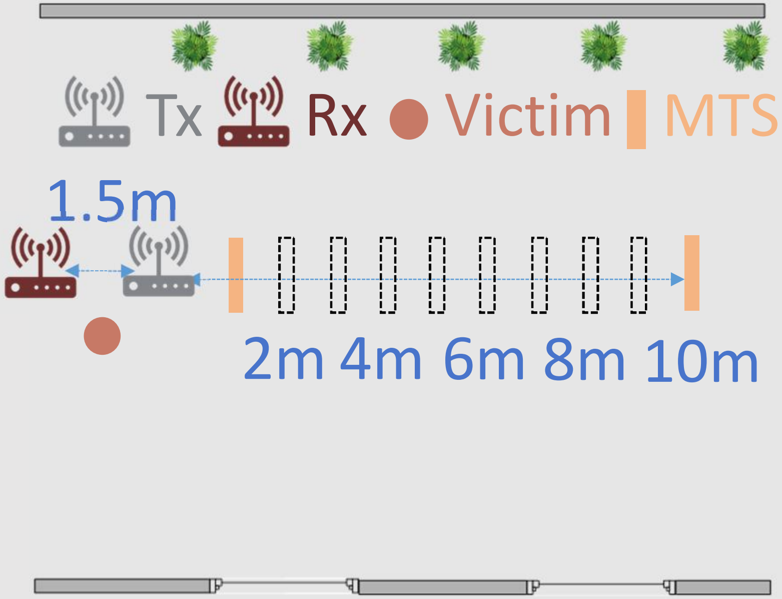




The average ASR on common materials is up to 80%

Evaluation

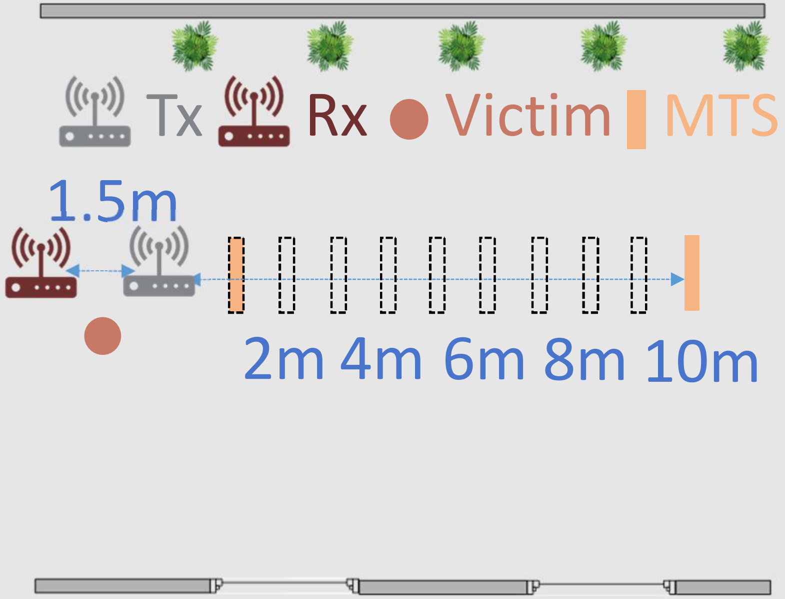
Performance under different obstacles.



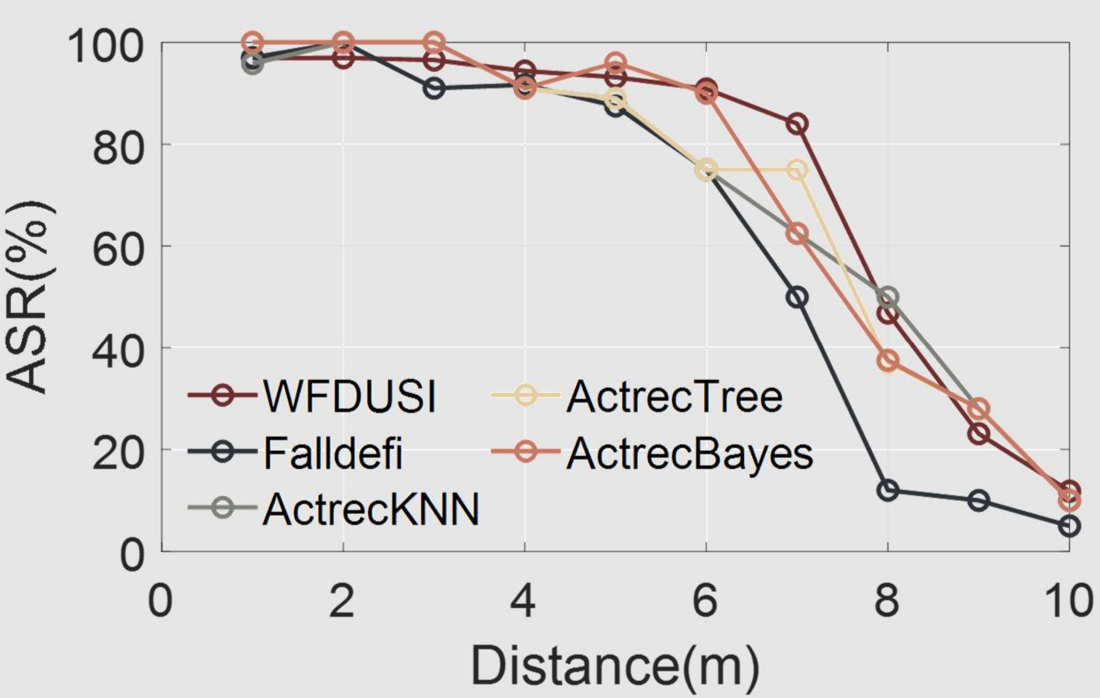
(a) The experiment layout

Evaluation

Performance under different obstacles.



(a) The experiment layout



(b) The ASR of different distances

Conclusions

1. RISiren designed a metasurface-assisted end-to-end black-box attack system against wireless sensing system with high stealthiness.

2. A novel attack scheme has been proposed to maximize the interference and generate human-like activity by carefully designing the approximation and optimization algorithm.

3. Only by changing the frequency-fit metasurface, RISiren can be easily generalized to other wireless sensing applications due to the nature of the metasurface being protocol-transparent.

4. Field study shows RISiren achieved attack success rate over 90% on average, and maintained robustness under different physical settings

THANK YOU