

# Challenges in Using AI/ML for Networking Research

Ram Durairajan

*Assistant Professor, CIS*

*University of Oregon*

Reza Rejaie

*Professor, CIS*

*University of Oregon*

Walter Willinger

*Chief Scientist*

*NIKSUN, Inc.*

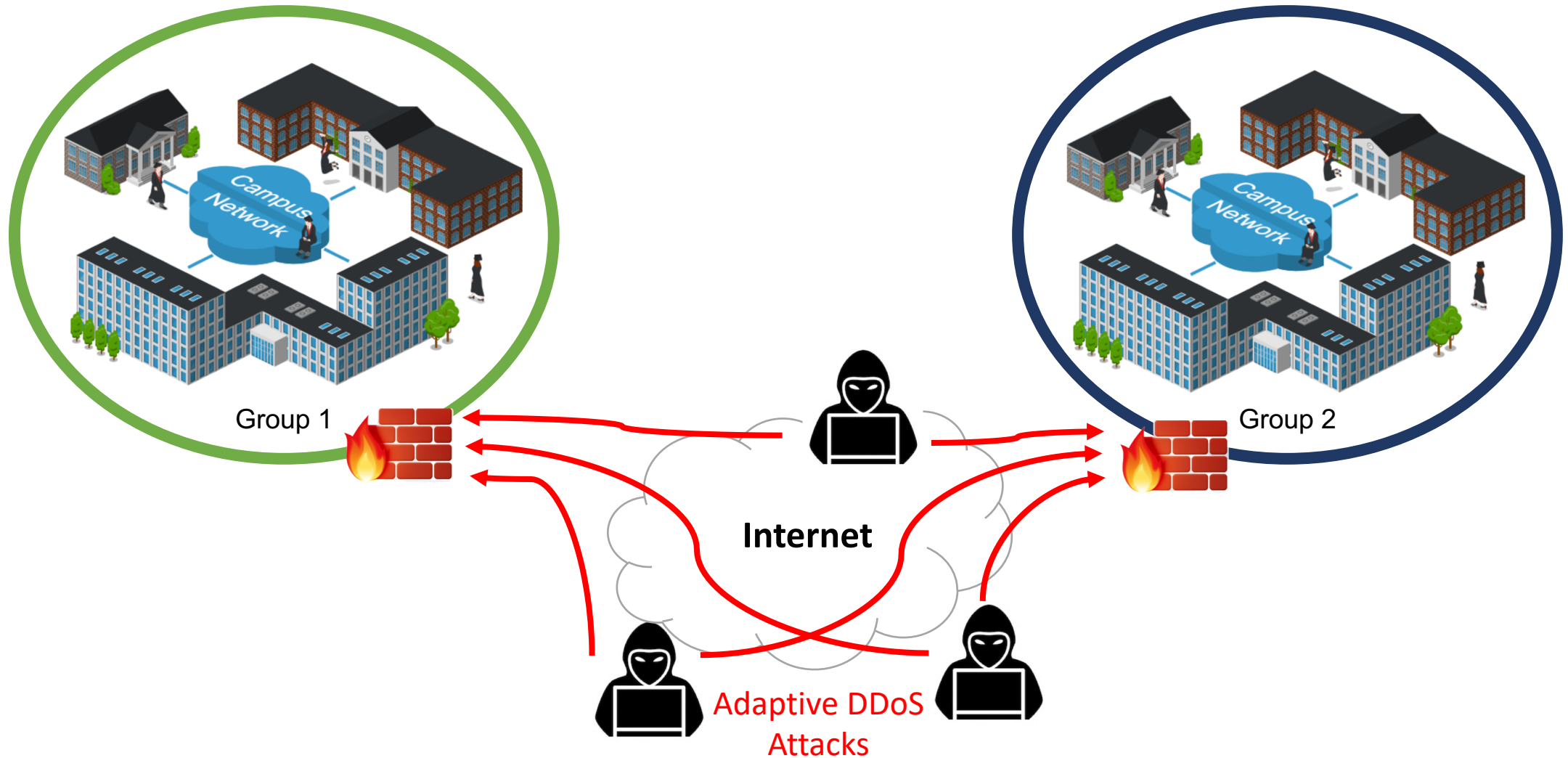
# Introduction

Enormous interest in ML for networking and cybersecurity

NO Data = NO ML!

Assuming the data problem is solved, the next problem awaits: **how to label if you must?**

# An Illustrative Example



# Challenges in using ML in networking

## Challenge 1: Lack of labeled networking data

Difficulty in labeling at scale

Current approaches are manual

Diversity of measurement data collected

Lack of agreement in community

What are the “k” typical features of a DDoS attack?

## Challenge 2: Privacy concerns in sharing the data

Sharing raw ~~or~~ labeled data

Sharing learning ~~models~~

Collaborate using  
ML in networking



# EMERGE: dEmocratize the use of ML for nEtwoRkinG rEsearch

Create high quality networking data labels using *data programming*

At scale



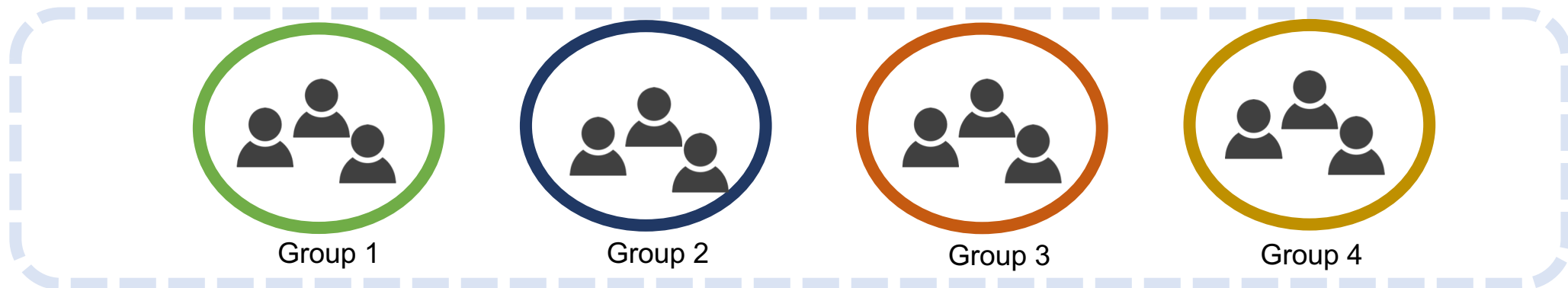
In programmable fashion



At low human labor cost



Promotes privacy-preserving collaborations: *only labeling functions are shared*



# Future Work

We are looking for partners to deploy EMERGE and do a pilot study

## Goals:

Identify events of interest and label them collaboratively

- \* Implication: standardization of features

Share labeling functions only – no data or models

- \* Implication: ensure privacy

Ultimately, learn about events/features from each other

- \* Implication: **robust AIOps for networking**

# Thank you!

## *Selected Papers:*

- [1] Y. Lavinia, R. Durairajan, R. Rejaie, and W. Willinger, “Challenges in Using ML for Networking Research: How to Label If You Must”, in proceedings of ACM SIGCOMM Workshop on Network Meets AI & ML (NetAI), August, 2020.
- [2] A. Muthukumar and R. Durairajan, “Denoising Internet Delay Measurements Using Weak Supervision”, in proceedings of IEEE ICMLA, December, 2019.

Source code: <https://gitlab.com/onrg/emerge>

We thank NSF for funding this project!

