A Machine-Learning Based Instruction Authentication Mechanism in Smart Home Networks

ICSML
International Cyber Security and Machine Learning
Academic & Professional Program

Weiming Bao & Kai Zhang
Team No. 11
1/8/2016

CONTENTS

1 Introduction and Motivation

2 Related work

3 Proposed mechanism

4 Preliminary experiments and result analysis

Introduction and Motivations

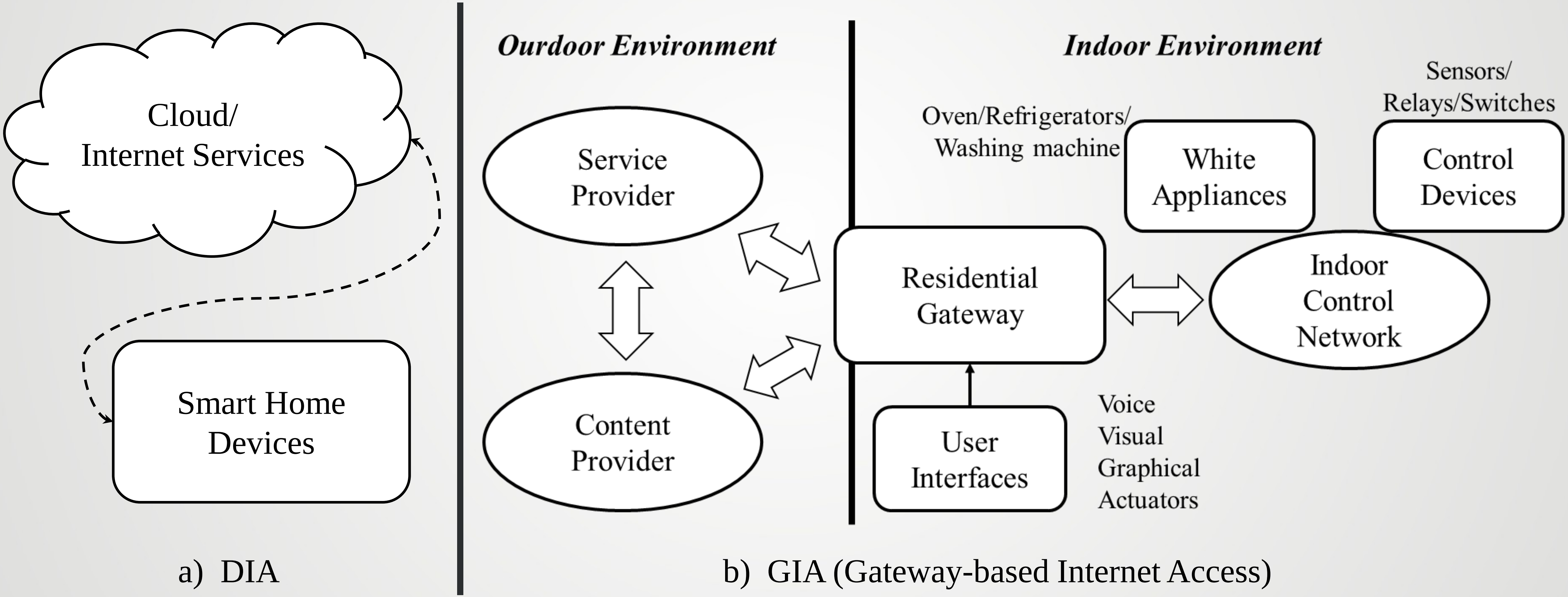
Smart Home

—— automatically and intelligently facilitate people’s everyday life and potentially provides additional comfort and even security

- Information / potential ability
- Internet access
- Constraints on hardware resources

Introduction and Motivations

Smart Home Network Architecture



Introduction and Motivations

Main vulnerabilities and attacks

Besides the traditional attacks, considering the CIAA of security ……

Extended Functionality Attacks

Unprofessional configuration

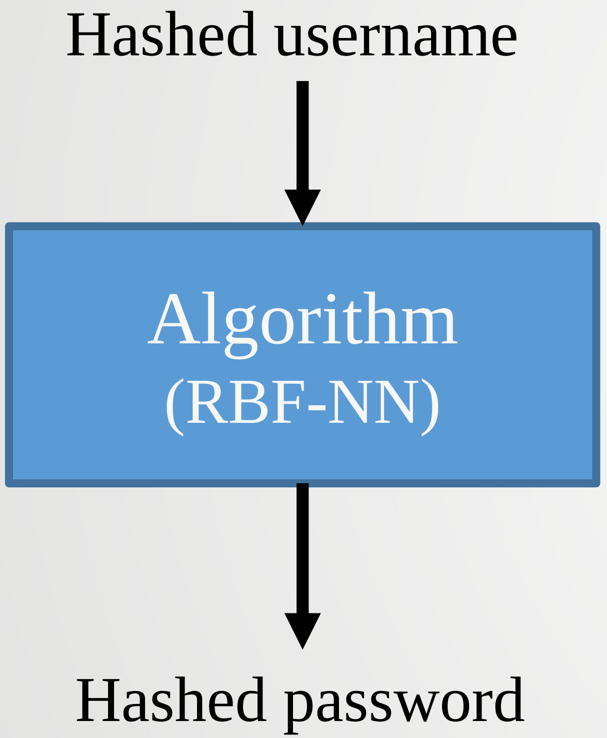
……

Authentication

Anomaly detection

Related works

Authentication using Neural Network in smart home networks

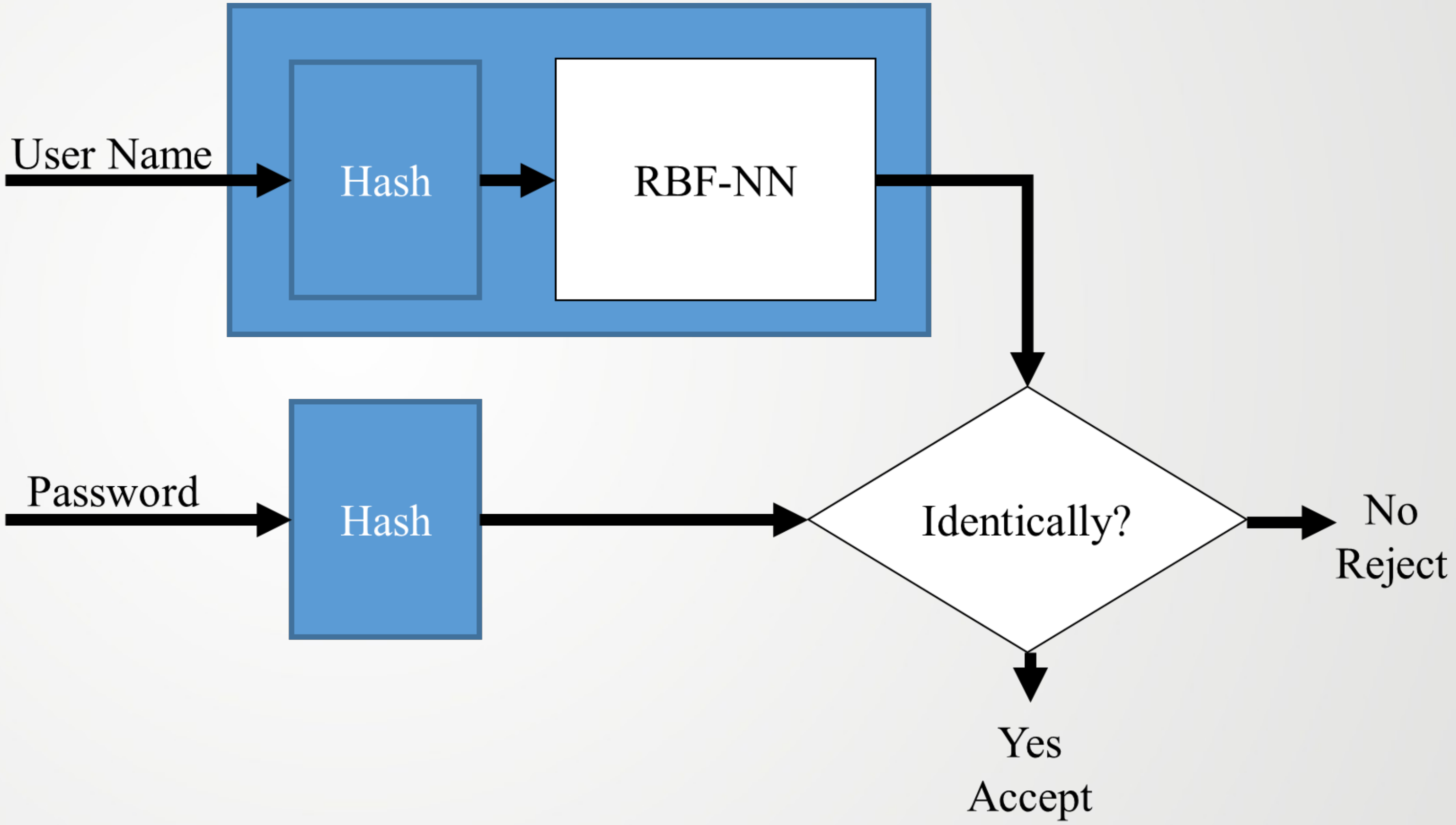


Hashed username

Algorithm

(RBF-NN)

Hashed password



User Name

Hash

RBF-NN

Password

Hash

Identically?

No

Reject

Yes

Accept

Shahbaz *et al*. User Authentication Using Neural Network in Smart Home Networks", *International Journal of Smart Home*, 2007

Related works

Activity and Anomaly Detection in Smart Home

Algorithm	Pros	Cons
Gaussian mixture model	Could relate two data attributes for activity classification	Reducing the matching times and eventually improve the detection efficiency

Hidden Markov model	Simple;
	Handling sequential data;
	Having temporal dependency structure;
	A statistical model that handles noisy data
		Need a full description of the big data;
		Requiring lots of trainings;
		Supervised Learning;
		Not fully capturing dependency structure of the data: a conditionally independent assumption

Artificial neural network	Being able to add new rules	Complex network architecture; Not understandable logic and rules behind the trained model;
Support vector machine	Provides a good out-of-sample generalization data; Linearly separable	Requiring 1-class CRF for anomaly detection when anomaly data instance is rare or unavailable

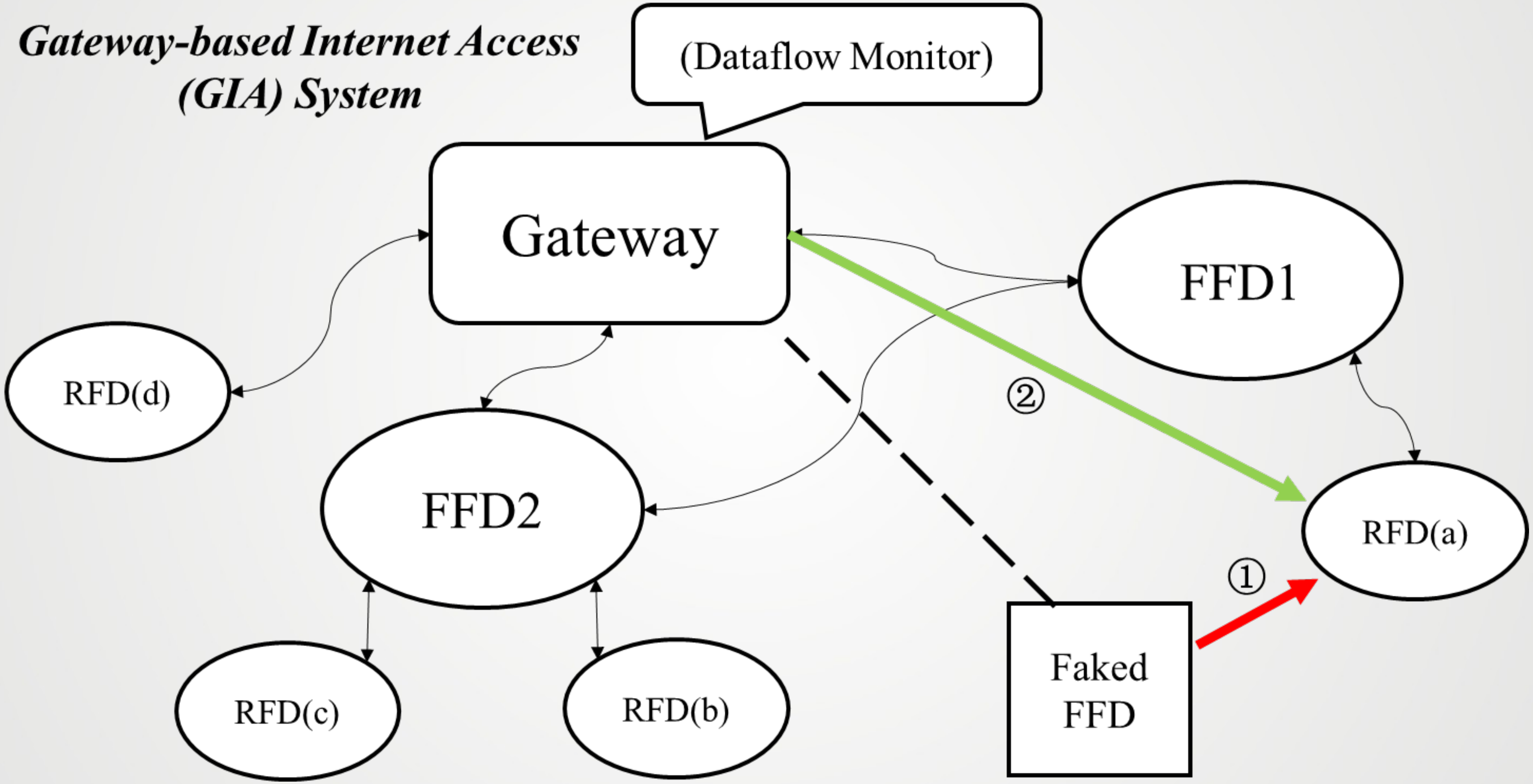
U.A.B.U.A. *et al*. Activity and Anomaly Detection in Smart Home: A Survey. *Next Generation Sensors and Systems*, 2015.

Proposed mechanism

Machine-Learning Based Instruction Authentication

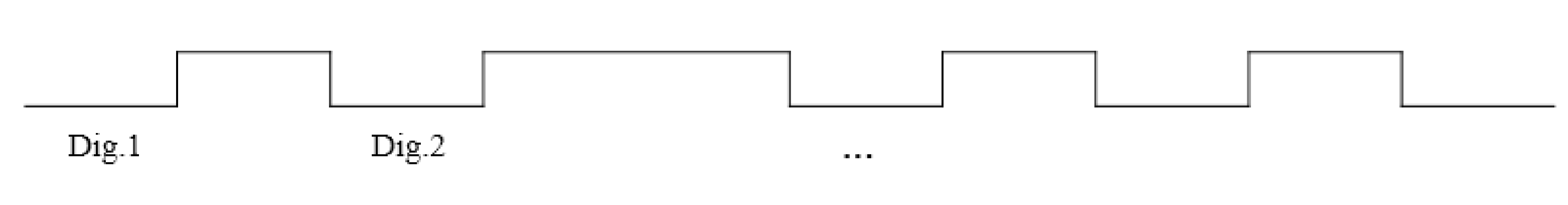
- Machine-learning based password verification
  - ◆ One-point verification
  - ◆ Password + Typing pattern (Username not needed)
  - ◆ Records and the model saved
- Anomaly detection based instructions legitimacy verification
  - ◆ GIA architecture system
  - ◆ Solution: block + re-authentication

Anomaly detection based instructions legitimacy verification



Data collection

The time series data of password and the typing pattern

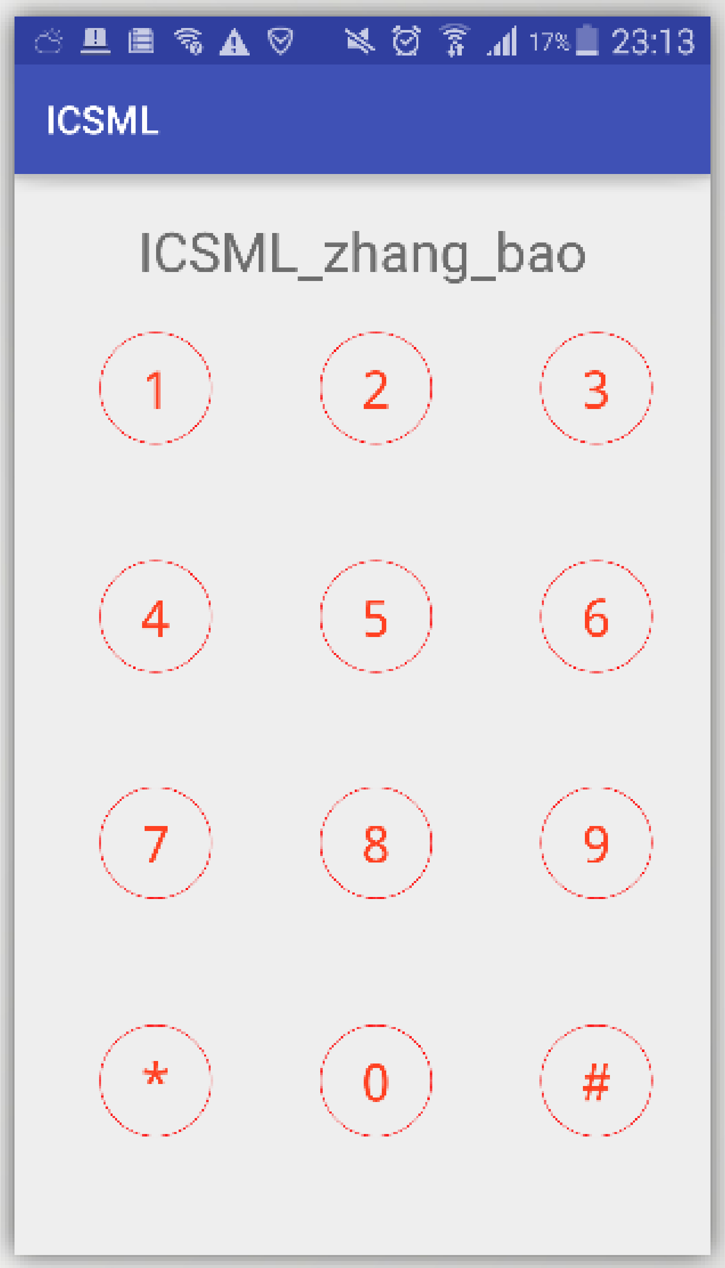


(touching / pressing / sweeping, at different pace)

Experiments

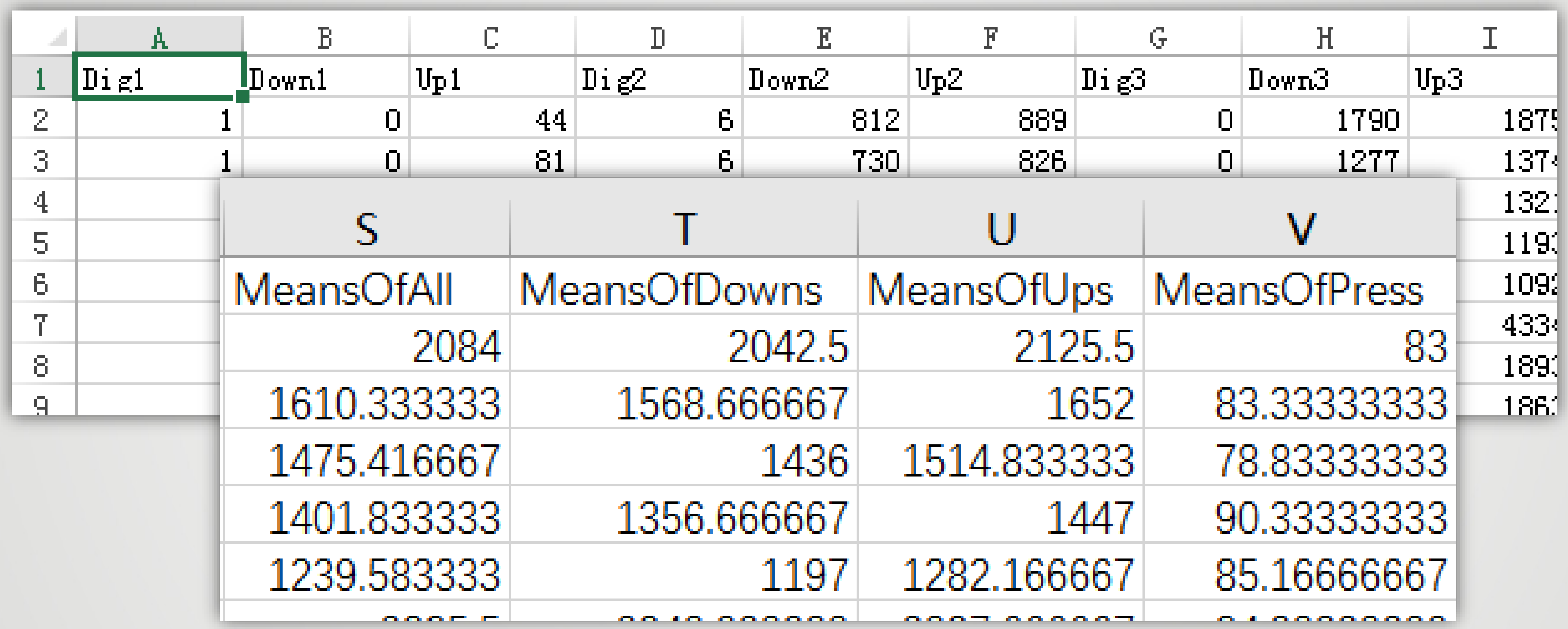
Data collection

The User Interface of the Android APP



Feature extraction

Preprocessed data: digits / press time / release time of the password / …



	A	B	C	D	E	F	G	H	I
1	Dig1	Down1	Up1	Dig2	Down2	Up2	Dig3	Down3	Up3
2	1	0	44	6	812	889	0	1790	1875
3	1	0	81	6	730	826	0	1277	1374
4									1321
5									1193
6									1092
7									4334
8									1893
9									1861

S	T	U	V
MeansOfAll	MeansOfDowns	MeansOfUps	MeansOfPress
2084	2042.5	2125.5	83
1610.333333	1568.666667	1652	83.33333333
1475.416667	1436	1514.833333	78.83333333
1401.833333	1356.666667	1447	90.33333333
1239.583333	1197	1282.166667	85.16666667

Experiments

Experiment sets

Normal non-featured PSW data + Intentionally **pattern** featured PSW data

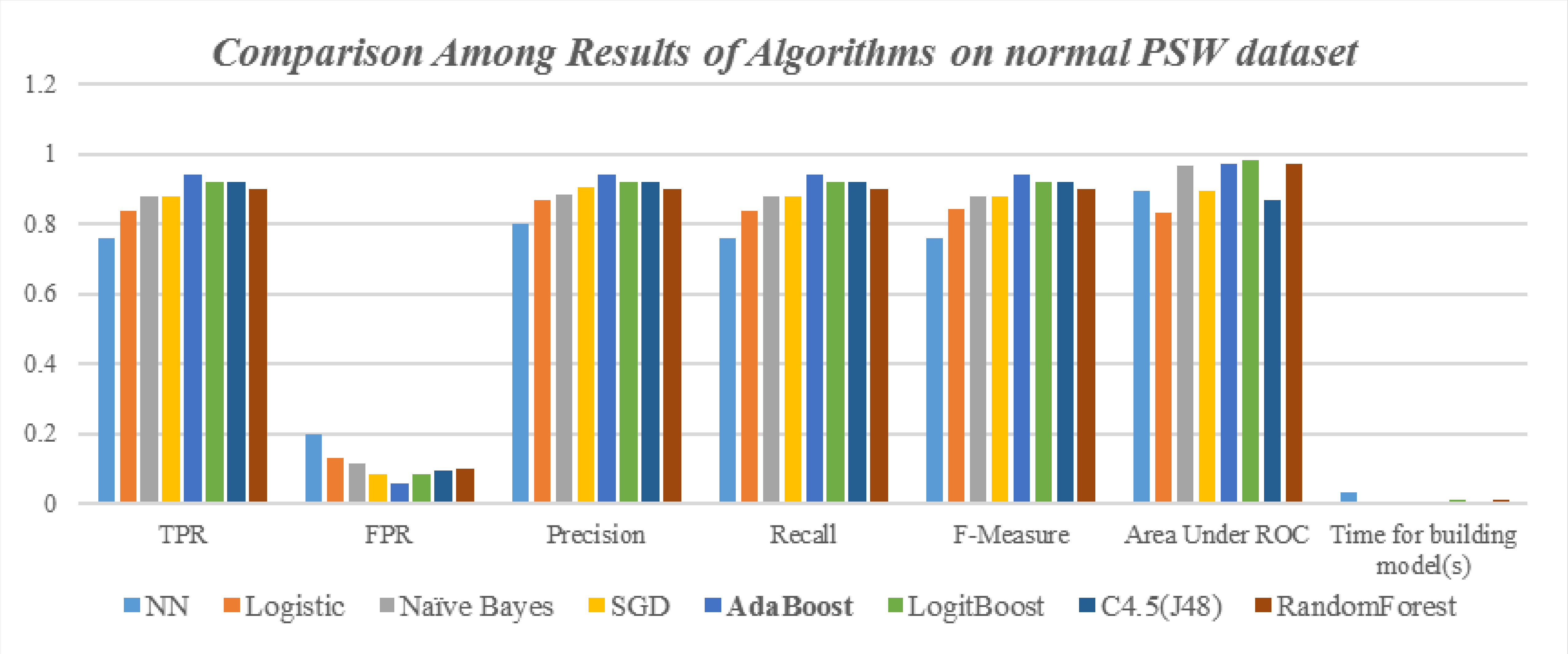
Do-wn1	Up1	Do-wn2	Up2	Do-wn3	Up3	Do-wn4	Up4	Do-wn5	Up5	Do-wn6	Standard Deviation (Downs)	Standard Deviation (Ups)	Category
104	586	82	563	95	608	83	541	61	529	85	13.23	28.86	Normal
84	651	84	484	107	451	75	408	63	683	64	14.89	110.58	Normal
85	528	74	650	107	506	86	428	86	538	87	9.78	71.341	Normal
731	639	161	934	64	351	43	319	65	450	76	244.82	227.02	Pattern
689	803	106	660	64	319	54	285	65	462	97	228.76	198.83	Pattern
690	494	86	550	75	341	42	310	51	451	65	233.82	90.83	Pattern

2 datasets

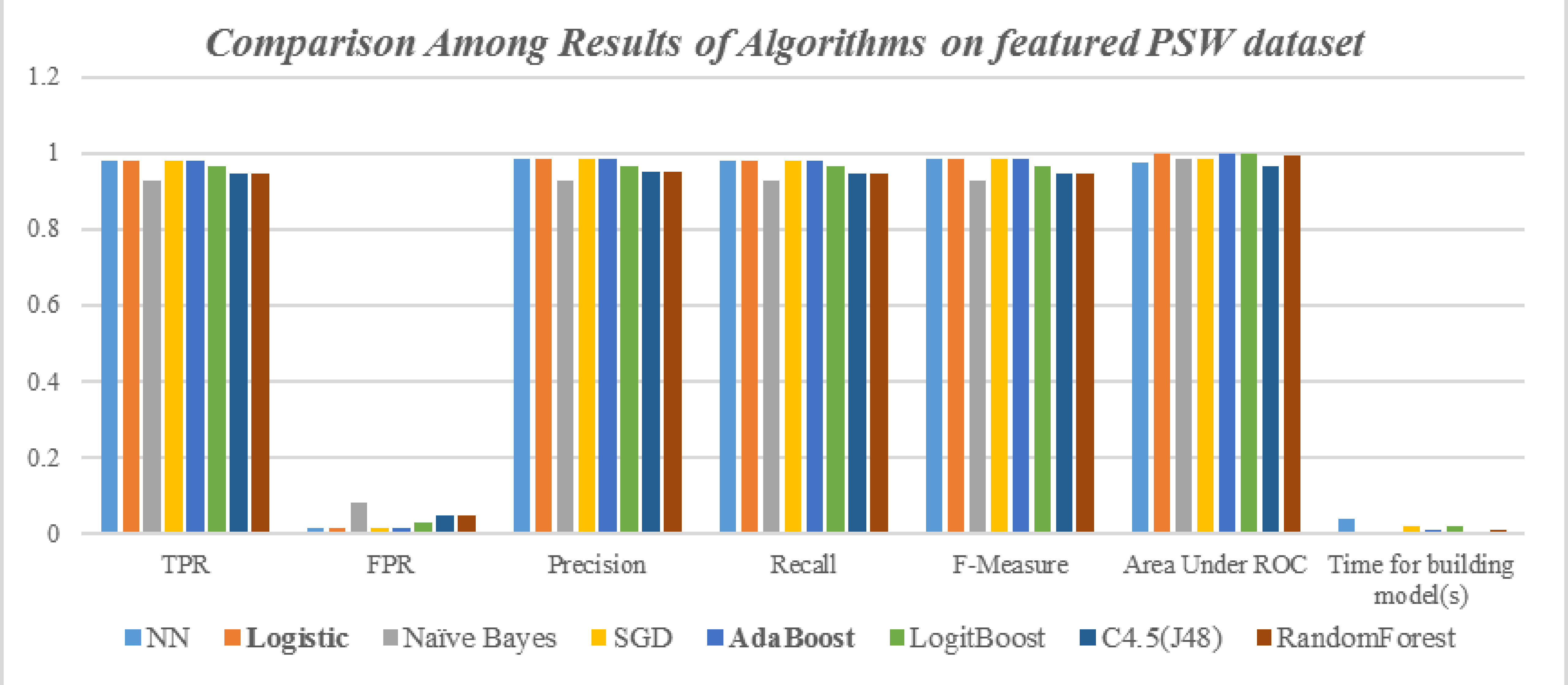
≈53 instances in each dataset

8 algorithms

10-fold cross validation/10 repetition


Comparison Among Results of Algorithms on normal PSW dataset

Algorithm	TPR	FPR	Precision	Recall	F-Measure	AUC	Training Time(s)
AdaBoost	0.94	0.057	0.941	0.94	0.94	0.975	<0.01

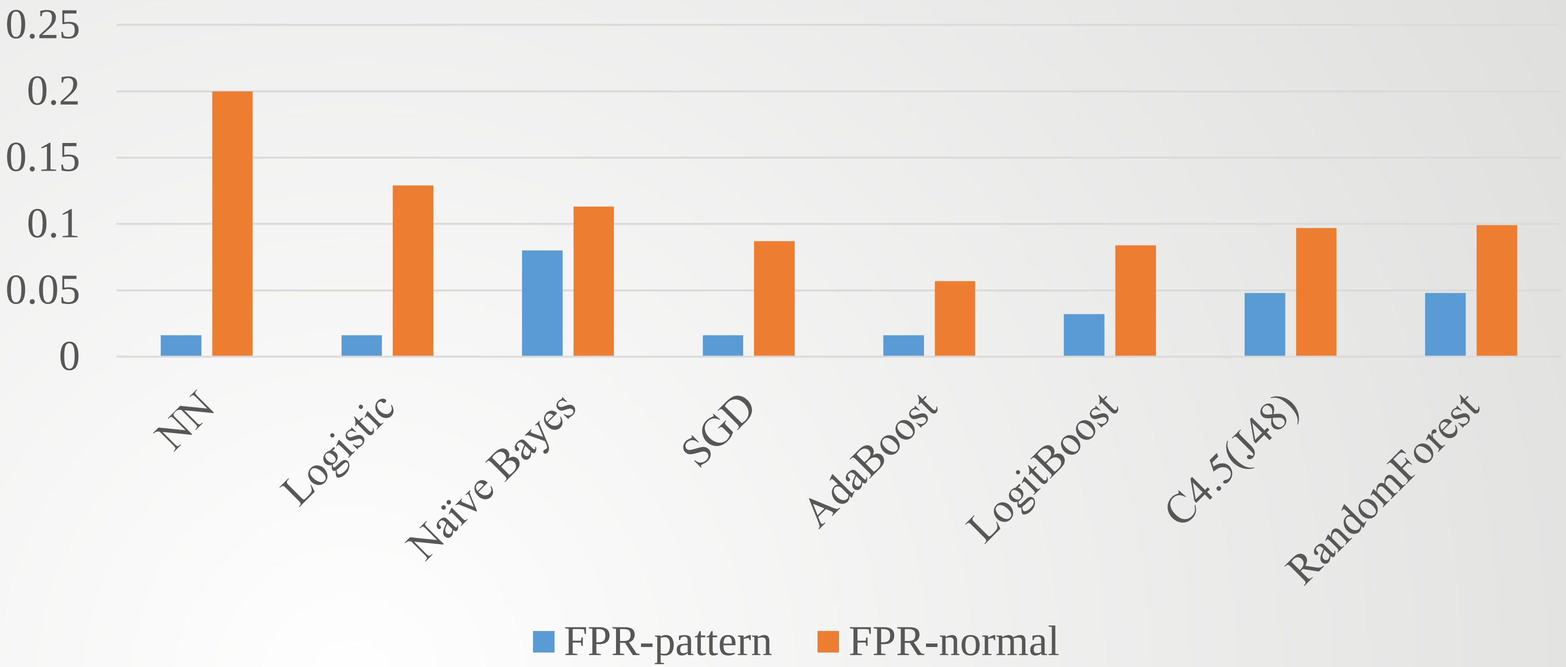


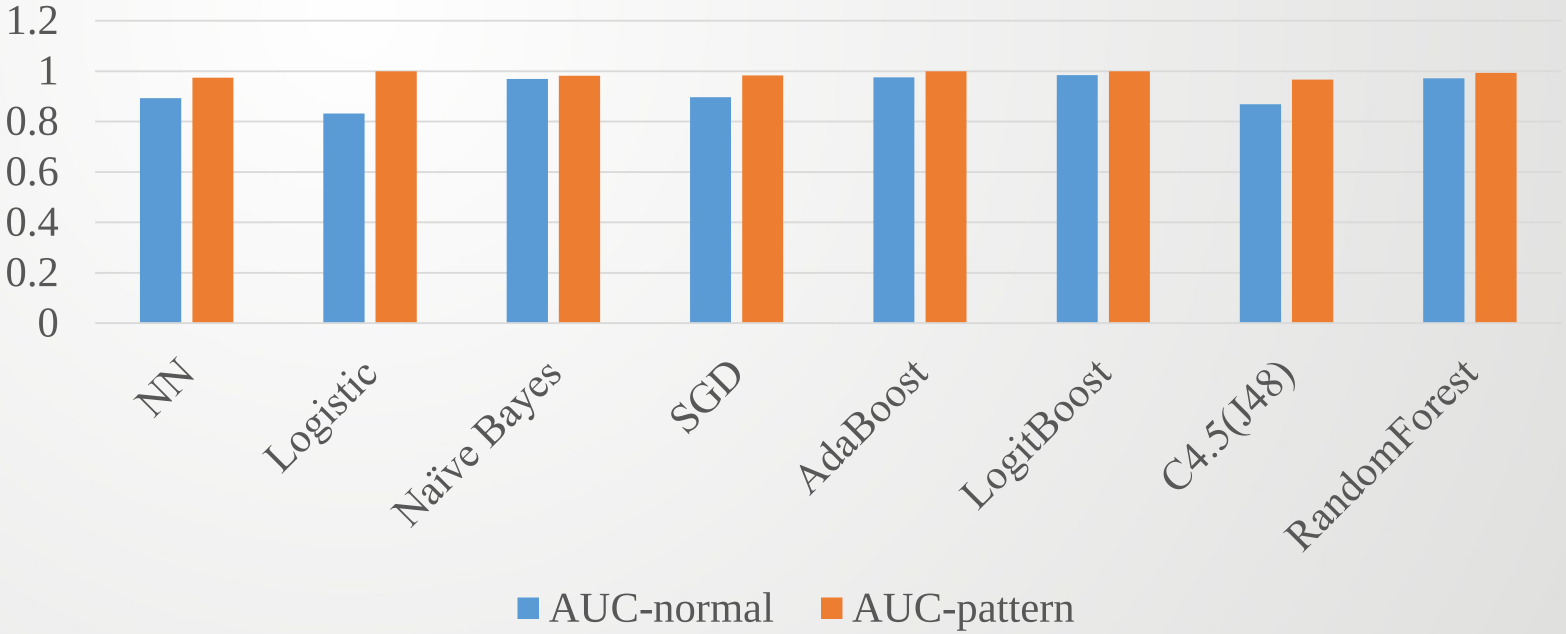
Algorithm	TPR	FPR	Precision	Recall	F-Measure	AUC	Training Time(s)
Logistic	0.981	0.016	0.982	0.981	0.982	1	<0.01

Results Analysis

Improvements

Algorithm	Enhancement on FPR
NN	92.0%
Logistic	87.6%
Naïve Bayes	29.2%
SGD	81.6%
AdaBoost	71.9%
LogitBoost	61.9%
C4.5(J48)	50.5%
RandomForest	51.5%





A Machine-Learning Based Instruction Authentication Mechanism in Smart Home Networks

Implement Dynamic Time Warping (DTW) algorithm

Experiment on real world Smart Home systems

Evaluate the performance and feasibility

A Machine-Learning Based Instruction Authentication Mechanism in Smart Home Networks

THANK YOU