Collaborative Security Attack Detection in Software-Defined Vehicular Networks

APNOMS 2017

Myeongsu Kim, Insun Jang, **Sukjin Choo**, Jungwoo Koo, and Sangheon Pack

Korea University

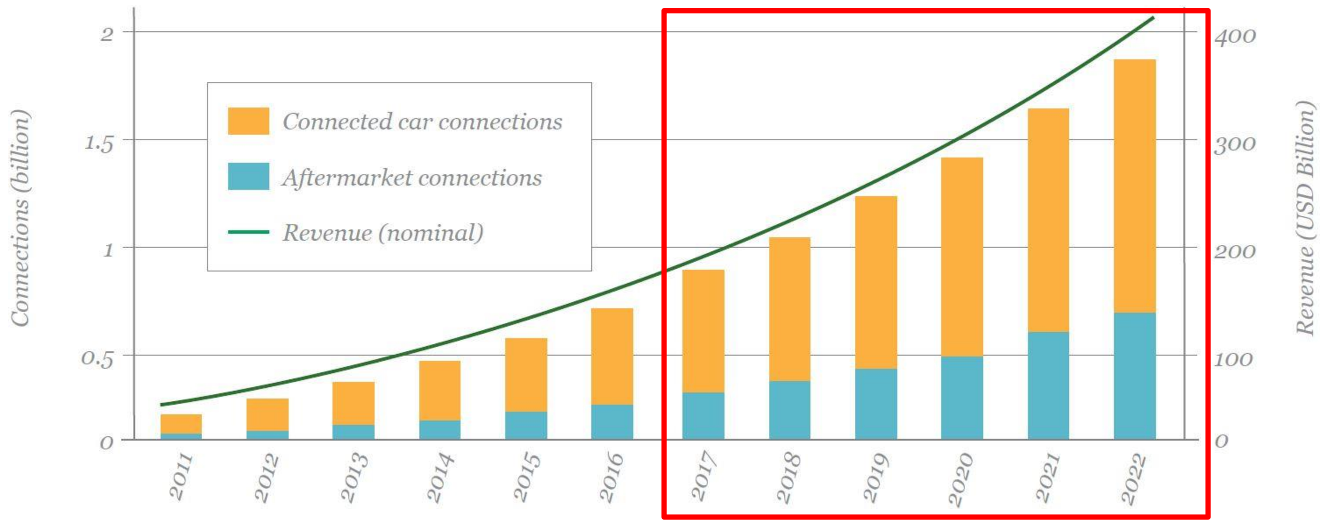
2017. 9. 27.

Contents

- Introduction
  - Software-defined Vehicular Cloud (SDVC)
- Collaborative security attack detection mechanism in software-defined vehicular networks
  - Motivation
  - Detection of attacks using multi-class SVM
- Simulation results
- Conclusion
- Reference

Introduction (1/2)

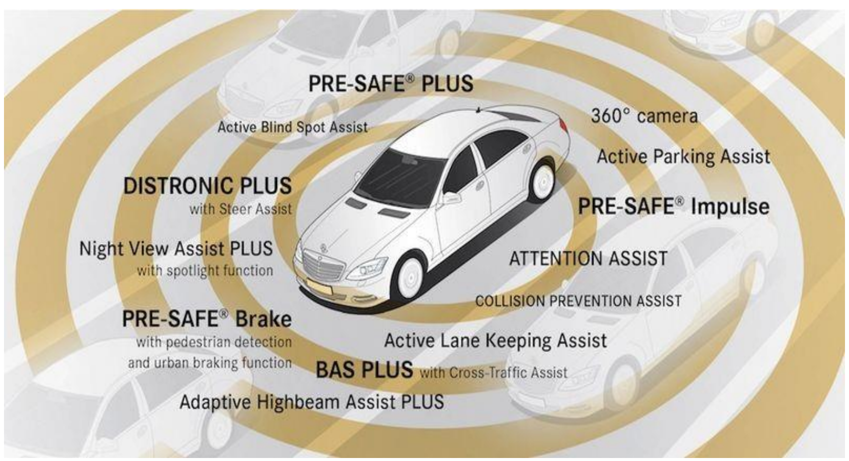
- The connected cars offer connectivity on wheels providing comfort and safety
  - Such an advanced technology enables the driver to connect with various online platforms or services
- The global connected car market has the potential to significantly boost revenues of car manufacturers



"Machine-to-machine connections and revenue in the automotive sector, 2011-2022"
[source: Machina Research, 2013]

Introduction (2/2)

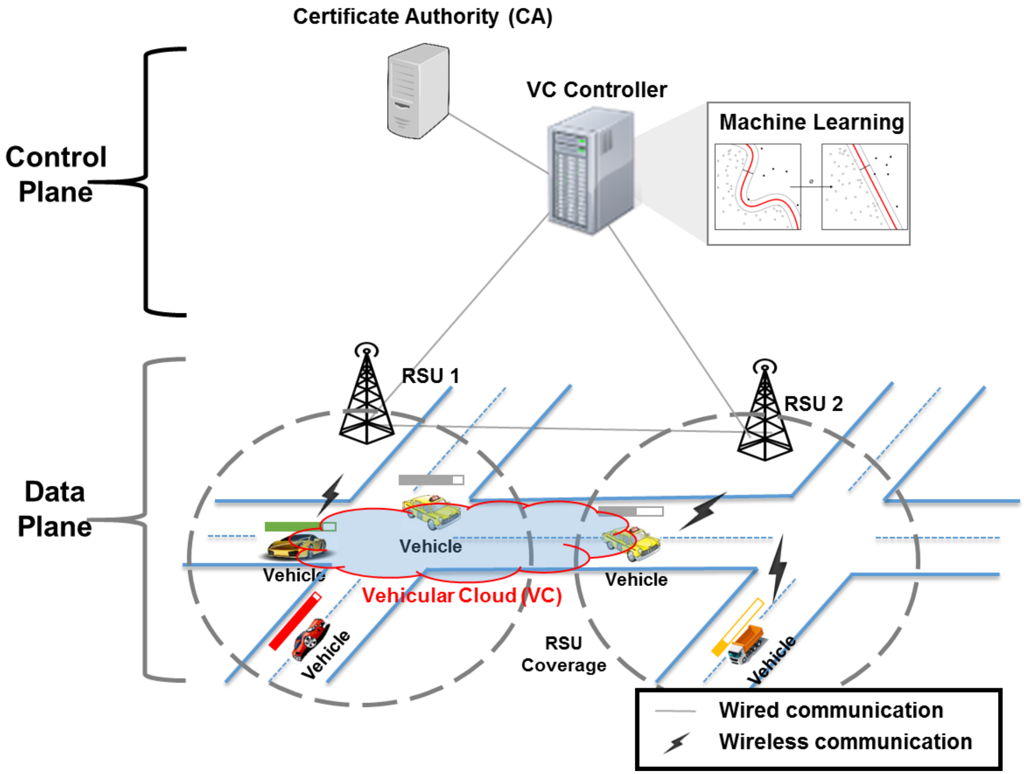
- In CES 2016, Qualcomm (with Audi) announced a Snapdragon 820 automotive processor for the connected cars
  - Qualcomm is providing the foundation for the next generation of infotainment platforms for automotive
  - E.g., Snapdragon LTE modem, IEEE 802.11ac, Bluetooth 4.1



[source: www.globalwindow.org, www.androidheadlines.com]

Software-defined Vehicular Cloud

- The resources of vehicles in VANETs are most likely not utilized (or under-utilized) for vehicular services
  - Computing, storage, and communication resource
- Software-defined Vehicular Cloud (SDVC) [1]



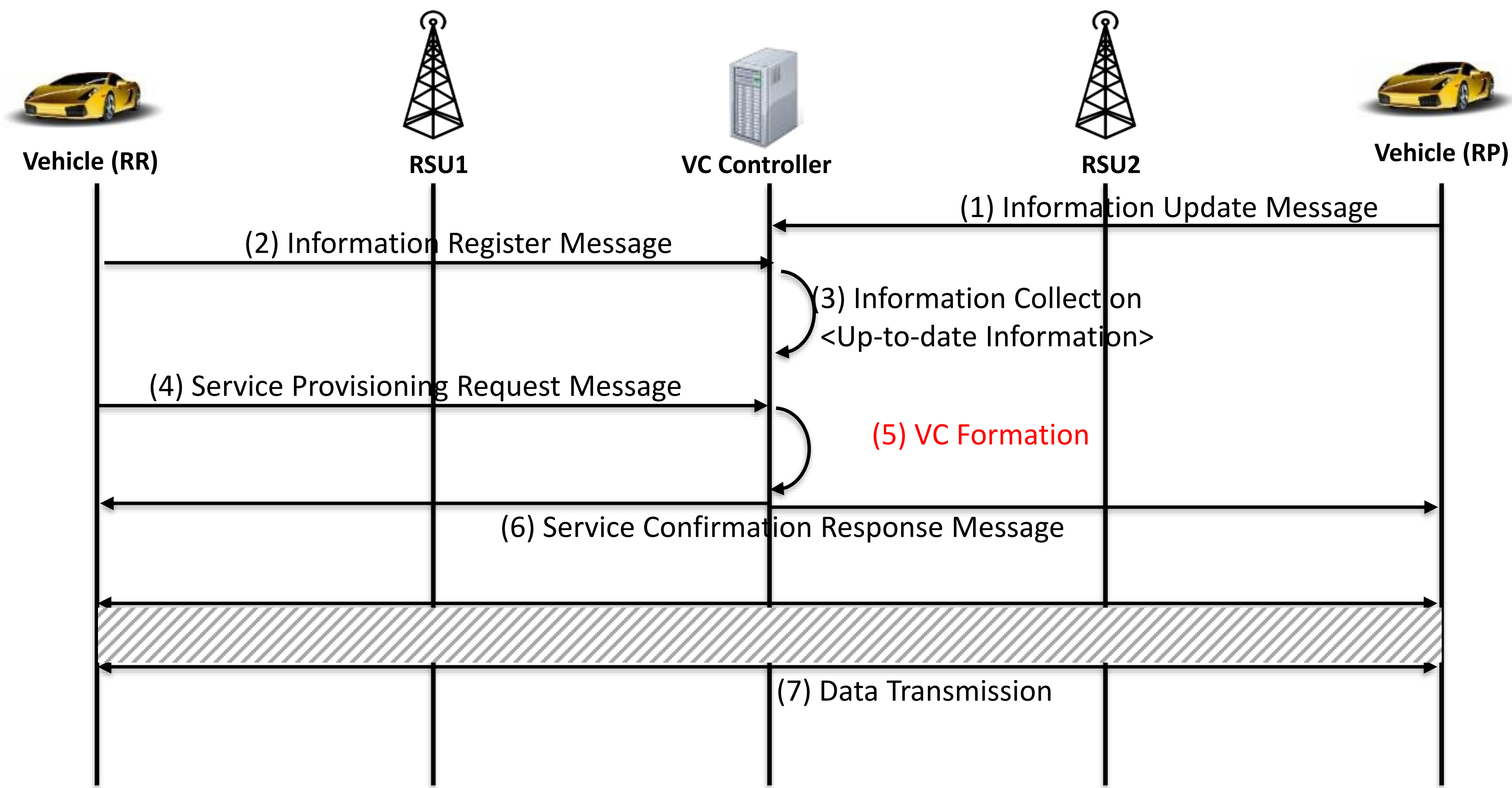
SDVC: Control plane

- Certificate authority (CA)
  - Assigns the public key and private key pairs along with the vehicle's certificate
- VC Controller
  - Collects global information of vehicles
    - E.g., vehicle ID, velocity, GPS location, and resource
  - Abstracts the vehicle's resources and maintains global view of vehicles
  - Performs resource distribution (i.e., VC formation) using V2X communications

SDVC: Data plane

- Vehicle
  - Registers local information of vehicles to the VC controller through the nearest RSU
    - Updates local information of vehicle to the VC controller periodically
  - Shares the resource via V2X communication
    - Type: Resource requester (RR), resource provider (RP)
- Road side unit (RSU)
  - Collects local information of vehicles
  - Forwarders information to the VC controller

SDVC: Operation



- Collaborative security attack detection mechanism in software-defined vehicular networks
  - Motivation
  - Detection of attacks using multi-class SVM

Collaborative security attack detection: Motivation (1/2)

- Security issues have been investigated in VANETs research [2]
- In traditional VANETs, a public key infrastructure (PKI) is commonly adopted by IEEE 1609.2 [3]
  - A certificate revocation list (CRL) is issued by the certificate authority (CA) periodically
  - There is no standard mechanism proposed for CRL
- The PKI can only ensure fundamental security requirements in VANETs
  - Authentication and message integrity

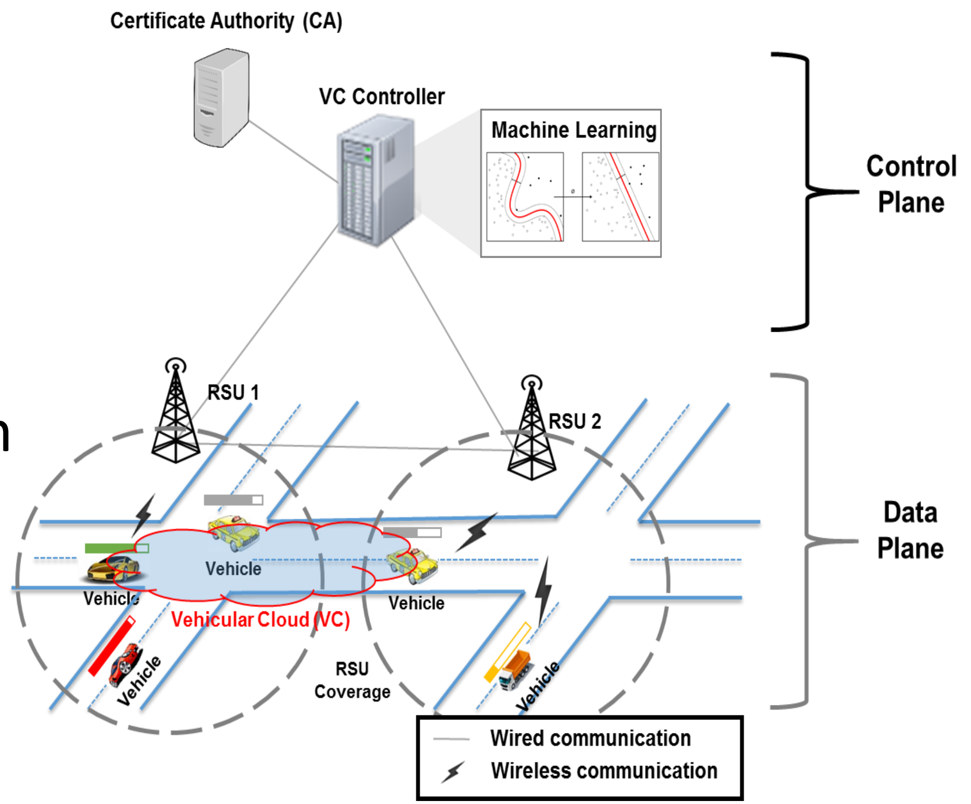
Collaborative security attack detection: Motivation (2/2)

- There are a number of attacks in VANETs [4][5]
  - Safety applications are very important in nature as these are directly related to drivers and their lives
  - The purpose of attacks is to create problem for drivers, and as a result services are not accessible
  - E.g., Sybil attack, denial of service (DoS) attack
- Attackers are moving and modifying their attack patterns continuously

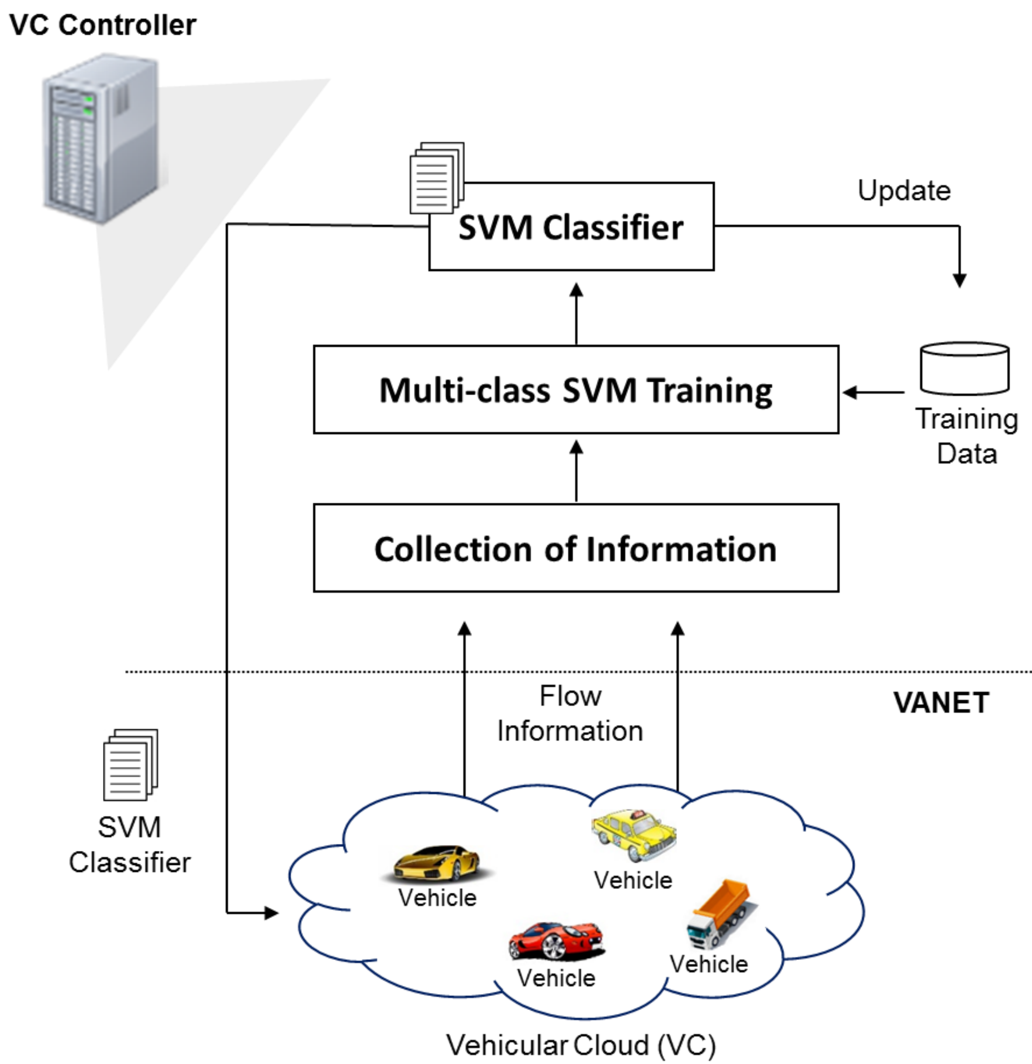
Collaborative security attack detection mechanism uses multi-class support vector machine (MC-SVM) to detect various types of attacks dynamically

Collaborative security attack detection: Overview

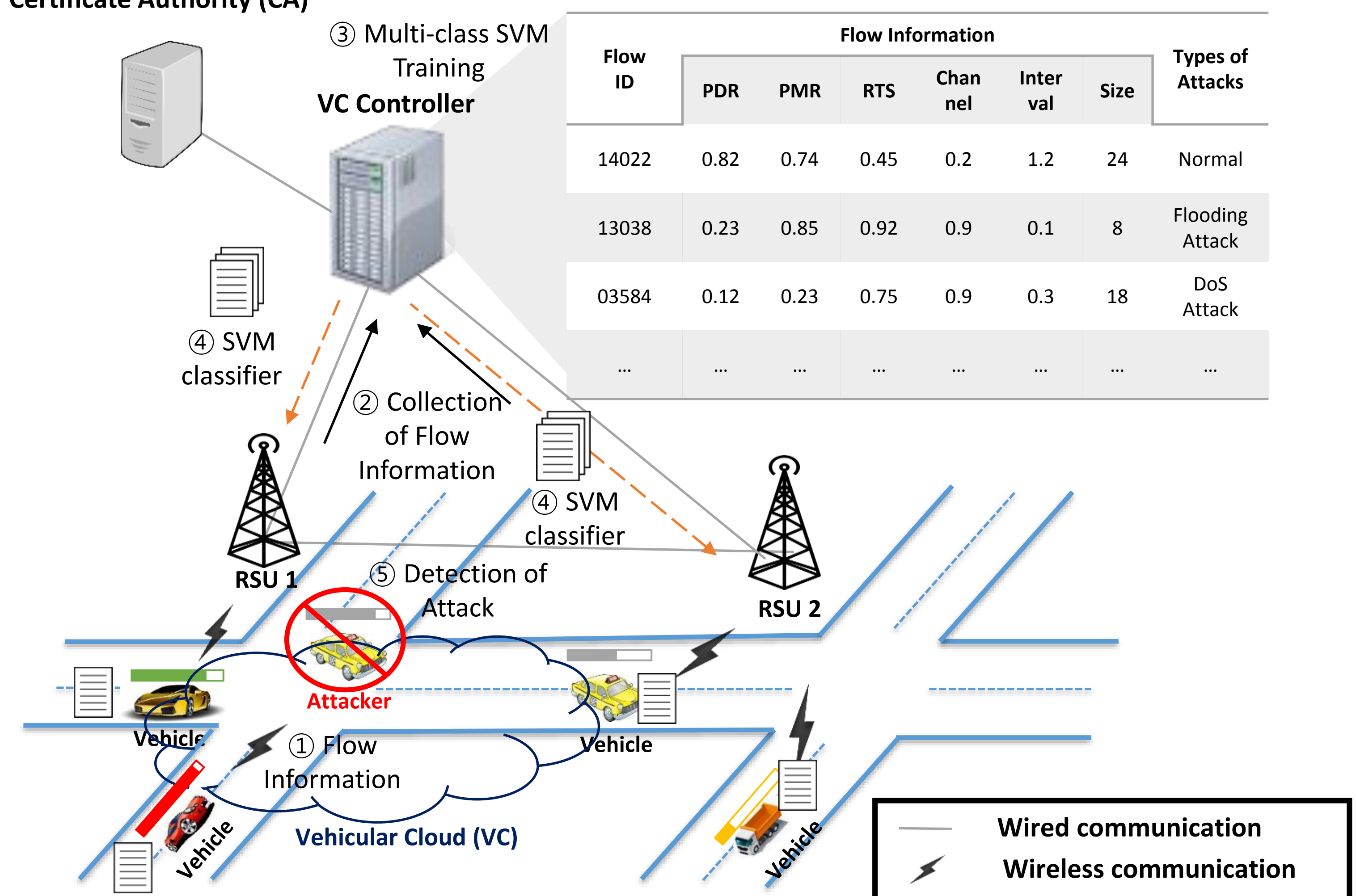
- Control plane
  - Certificate authority (CA)
    - Issuing the certificate
  - VC controller
    - Information collection
    - VC formation
    - Generating pseudonym
    - Conducting multi-class SVM
- Data plane
  - Road segment unit (RSU)
  - Vehicle



Collaborative security attack detection: Operation



Detection of attacks using MC-SVM: Example

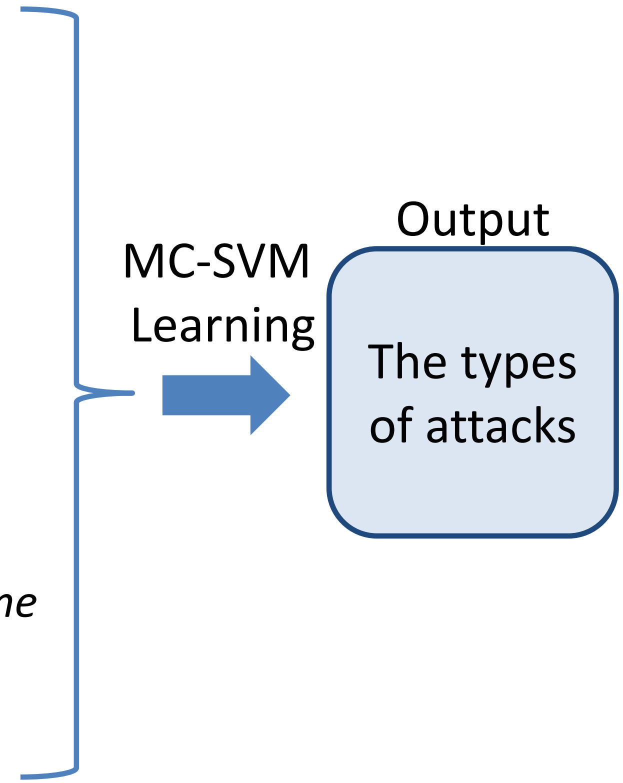


Multi-class SVM Training Data

Flow ID	Flow Information						Types of Attacks
	PDR	PMR	RTS	Chan nel	Inter val	Size	
14022	0.82	0.74	0.45	0.2	1.2	24	Normal
13038	0.23	0.85	0.92	0.9	0.1	8	Flooding Attack
03584	0.12	0.23	0.75	0.9	0.3	18	DoS Attack
…	…	…	…	…	…	…	…

Detection of attacks using MC-SVM: Modeling

- Multi-class SVM features
  - Packet drop rate (PDR)
    - $PDR = \frac{The\ Number\ of\ Packets\ Dropped}{The\ Total\ Number\ of\ Packets\ Transmitted}$
  - Packet modification rate (PMR)
    - $PMR = \frac{The\ Number\ of\ Packets\ Modified}{The\ Total\ Number\ of\ Incoming\ Packets}$
  - RTS flooding rate
    - *IEEE 802.11p RTS packet*
  - Wireless channel status [0, 1]
    - *Busy status of channel in a specific period of time*
  - Packet interval, packet size
    - *Average packet interval and size in the flow*

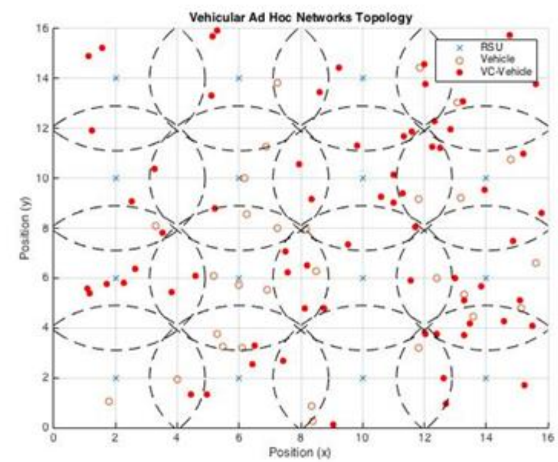


Simulation results: Topology

- MC-SVM simulator based on Matlab 2015a
  - Dataset: KDD Cup 1999 (by MIT Lincoln Labs) *
    - The objective is to survey and evaluate research in IDS
    - Attacks: DoS, Probing, R2L, U2R + Normal (# 86,678 dump (10%))
  - Comparison scheme
    - SVM-Nearest Neighbor, SVM-Individual
  - Simulation parameters

Parameter	Value
Simulation area	1,600 m × 1,600 m
The number of RSUs	16
Transmission range of RSU	300 m
The number of vehicles	10, 20, 30, 40, 50
Transmission range of vehicle	130 m

Random Generation

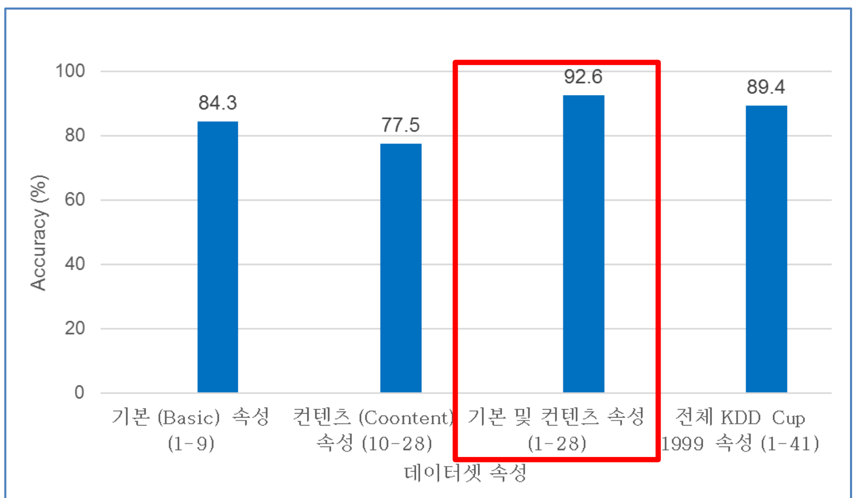


[*] http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

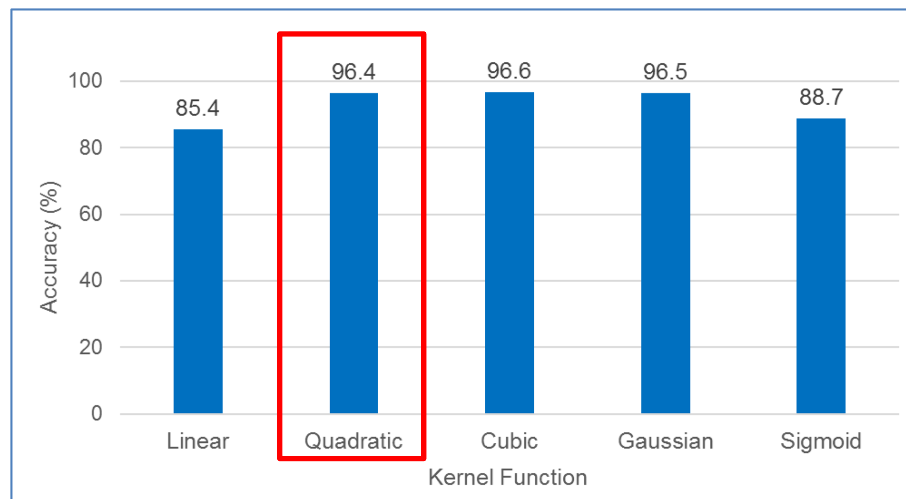
Simulation results: KDD Cup 1999 dataset features

- KDD Cup 1999 dataset features
  - Basic features (1-9) ➡ [DoS, Probing attack]
    - duration, protocol, service, flag, src_byte, dst_byte, land, wrong_fragment, urgen
  - Content features (10-28) ➡ [R2L, U2L attack]
    - count, srv_count, serror_rate, srv_serror_rate, …

KDD Cup 1999 features

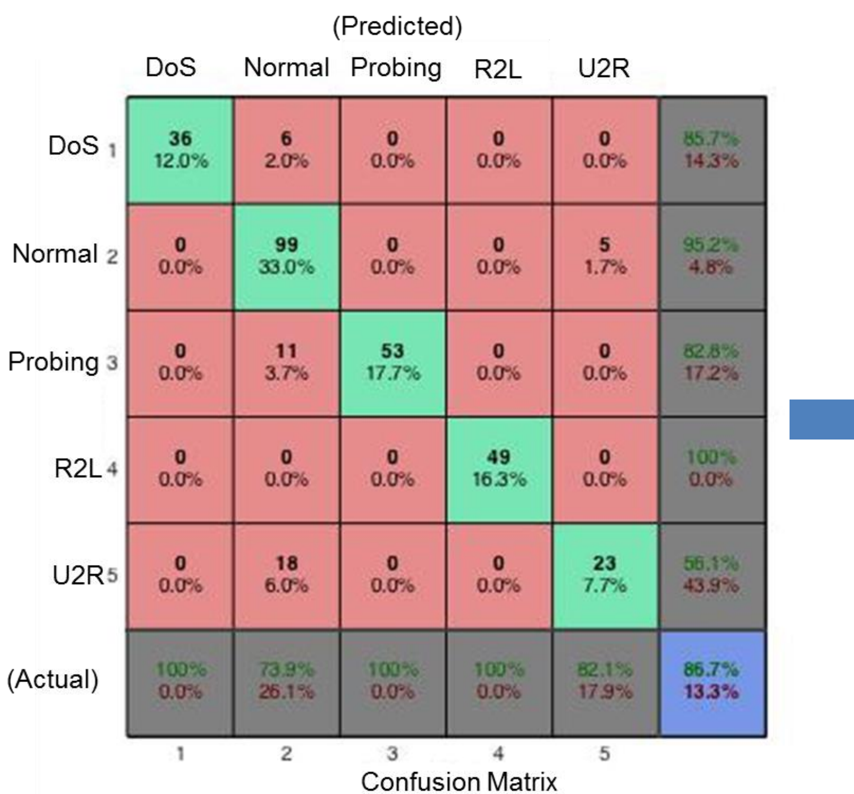


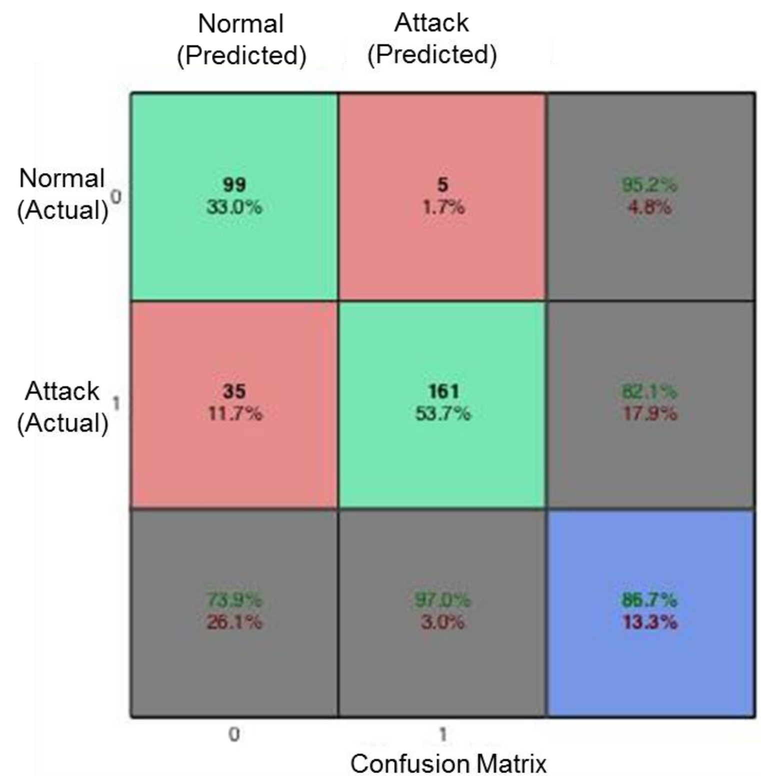
MC-SVM kernel function



Simulation results: Confusion matrix

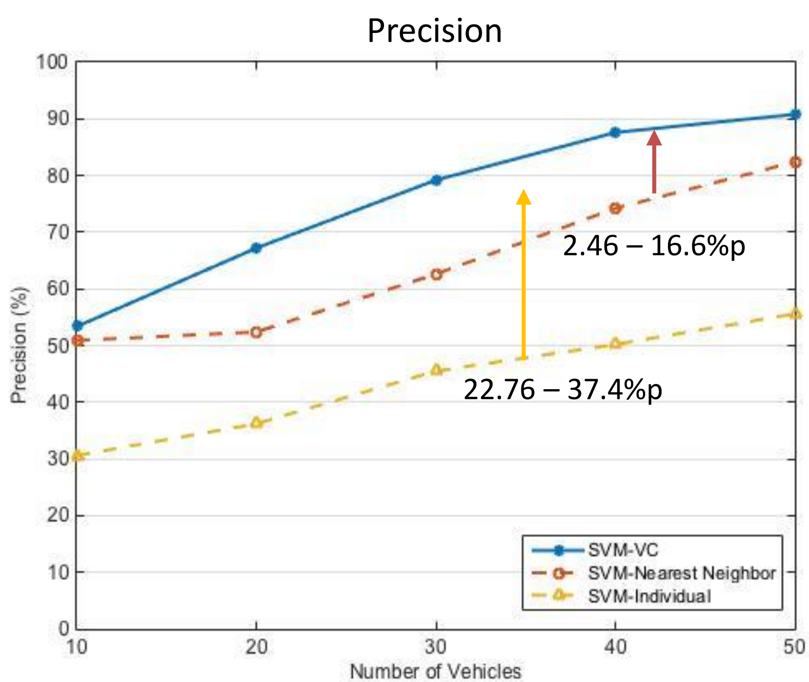
- Confusion matrix
  - Test dataset: # 300

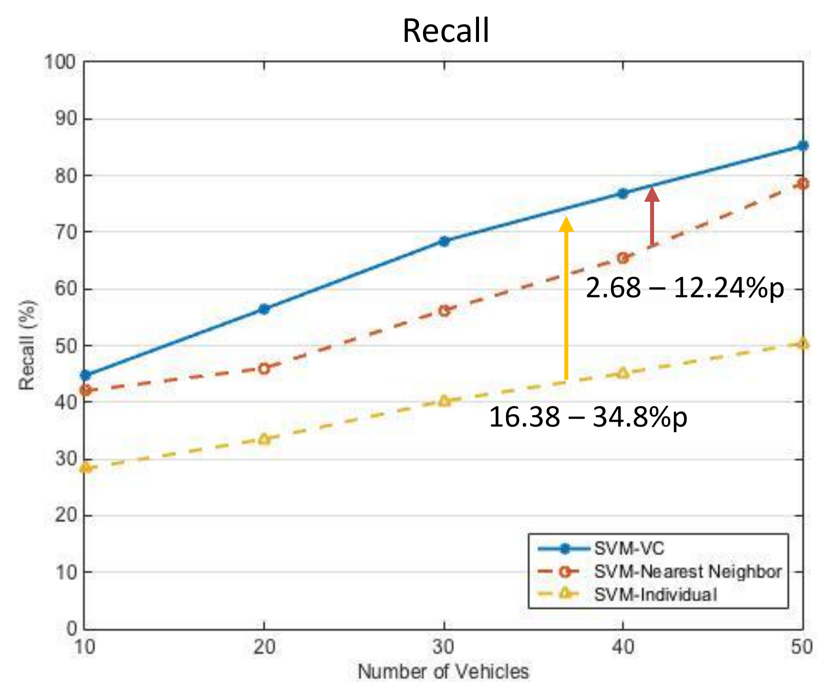




Simulation results:
Effect of vehicle density (1/2)

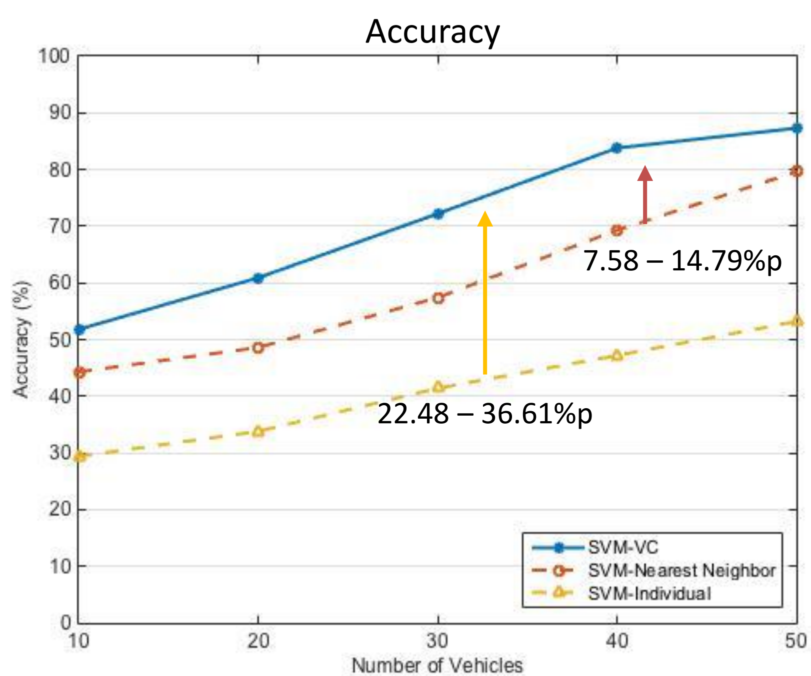
- The number of vehicles: [10, 20, 30, 40, 50]
  - MC-SVM dataset: #30,000 (Learning), # 20,000 (Test)
  - Vehicle: Random (# 100 – 1,000)





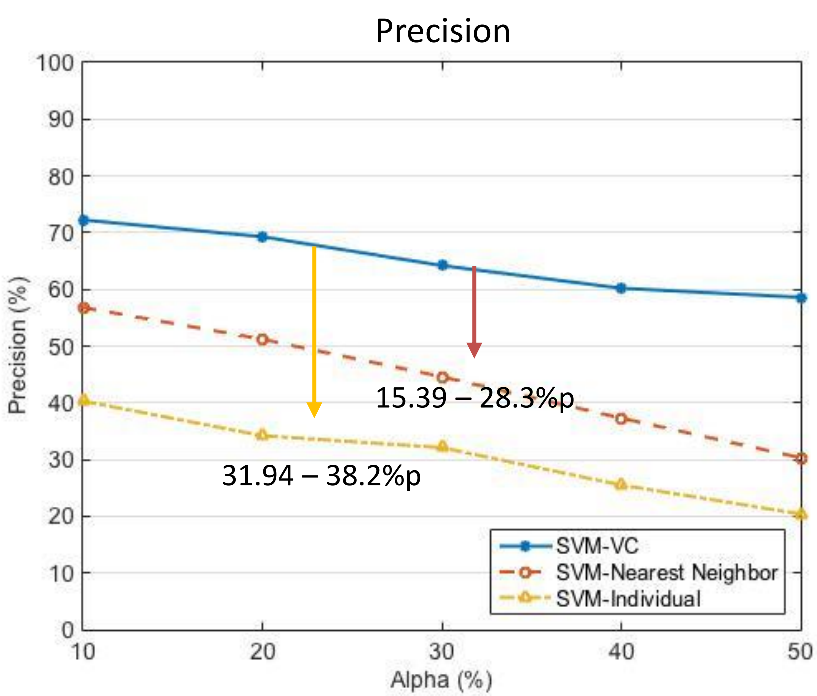
Simulation results: Effect of vehicle density (2/2)

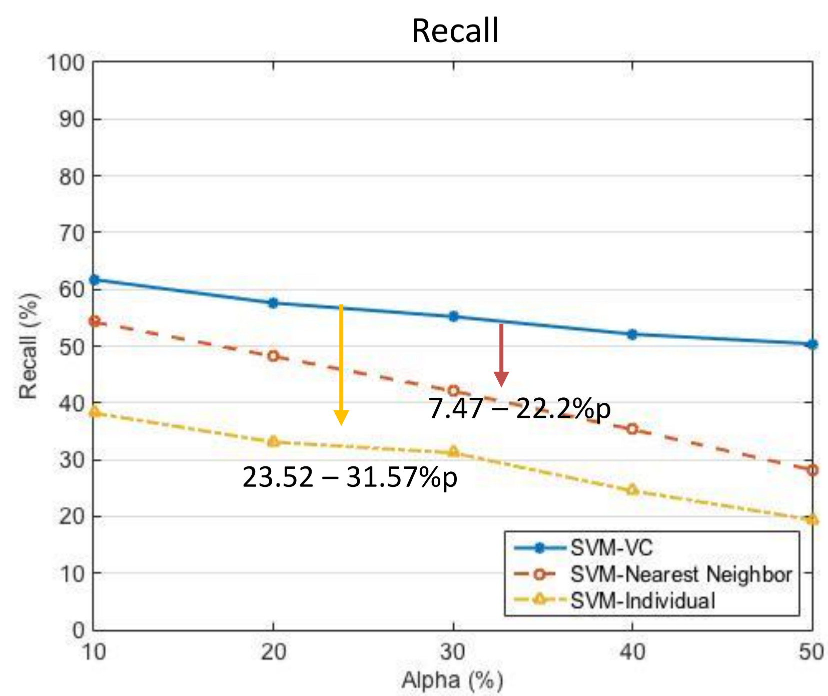
- The number of vehicles: [10, 20, 30, 40, 50]
  - MC-SVM dataset: #30,000 (Learning), # 20,000 (Test)
  - Vehicle: Random (# 100 – 1,000)



Simulation results: Effect of alpha (1/2)

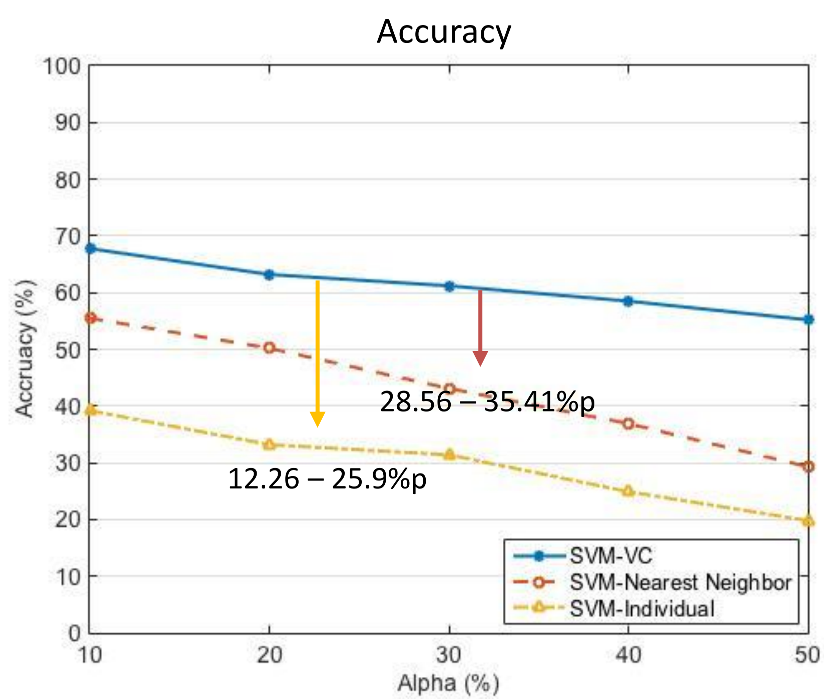
- The variation of alpha (%): [10, 20, 30, 40, 50]
  - MC-SVM dataset: #30,000 (Learning), # 20,000 (Test)
  - Vehicle: Random (# 100 – 1,000)





Simulation results: Effect of alpha (2/2)

- The variation of alpha (%): [10, 20, 30, 40, 50]
  - MC-SVM dataset: #30,000 (Learning), # 20,000 (Test)
  - Vehicle: Random (# 100 – 1,000)



Conclusion

- We proposed collaborative security attack detection mechanism in software-defined vehicular networks
  - we use multi-class support vector machine (MC-SVM) to detect various types of attacks
  - The simulation results show that the proposed mechanism achieves a good performance to detect the types of attacks
    - High precision, recall, and accuracy
  - In our future works, we will extend MC-SVM model to minimize the network bandwidth usage

Reference

[1] S. Choo, I. Jang, M. Kim, and S. Pack, “The Software-Defined Vehicular Cloud: A New Level of Sharing the Road,” *IEEE Vehicular Technology Magazine (VTM),* vol. 12, no.2, pp. 78-88, June 2017.

[2] F. Qu, Z. Wu, F. -Y. Wang, and W. Cho, “A Security and Privacy Review of VANETs,” *IEEE Transactions On Intelligent Transportation Systems*, vol. 16, no. 6, pp. 2985-2996, Dec. 2015.

[3] N. Tiwari, “On the Security of Pairing-free Certificateless Digital Signature Schemes Using ECC,” *ICT Express*, vol. 1, no. 2, pp. 94-95, Sept. 2015.

[4] M. Azees, P. Vijayakumar, and L. J. Deborah, “Comprehensive Survey of Security Service in Vehicular Ad-hoc Networks,” *IET Intelligent Transport Systems*, vol. 10, no. 6, pp. 379-388, Aug. 2016.

[5] L. Barish, D. Shehada, E. Salahat, and C. Y. Yeun, “Recent Advances in VANET Security: A Survey,” in *Proc. IEEE Vehicular Technology Conference (VTC) Fall*, 2015.

[6] W. Li, A. Joshi, and T. Finin, “SVM-CASE: An SVM-based Context Aware Security Framework for Vehicular Ad-hoc Networks,” in *Proc. IEEE VTC Fall*, Sept. 2015.

Q & A

Backup

MC-SVM: Modeling

- Let, $\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_n, y_n)\}$,
  $$where\ x_i \in R^D, y_i \in 0,1,2, \dots, m, i = 1,2, \dots, n.$$
- The decision boundary should be classify all points correctly

$$y_i(w^T x_i + b) \geq 1, \quad 1 \leq i \leq n.$$

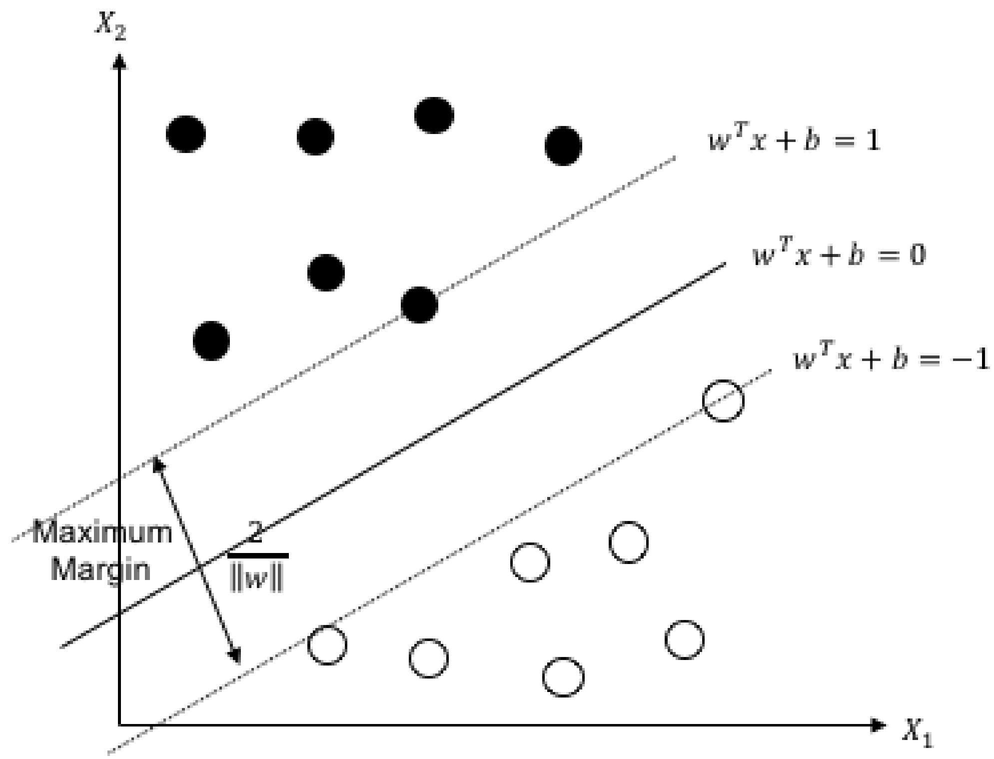
- The decision boundary can be found by solving the following constrained optimization problem

$$\min_{(w,b)} \frac{1}{2} \|w\|^2$$

$$subject\ \ to\ \ y_i(w^T x_i + b) \geq 1, \quad 1 \leq i \leq \mathrm{n}.$$

MC-SVM: Modeling

- The decision boundary should be as far away from the data of both as classes possible
  - The goal is to maximize the margin, *m*



MC-SVM: Modeling

- Converts to convex optimization problem using slack variable,

$$\min_{(w,b)} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i$$

$$subject\ \ to\ \ \ y_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, 1 \leq i \leq \mathrm{n}.$$

- Transforms dual problem using Lagrange multiplier formula,

$$\max_{(\alpha)} L(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i\, \alpha_j y_i y_j\, K(x_i, y_j)$$

$$subject\ \ to \sum_{i=1}^{n} \alpha_i\, y_i = 0, \qquad 0 \leq \alpha_i \leq C, 1 \leq i \leq \mathrm{n}.$$

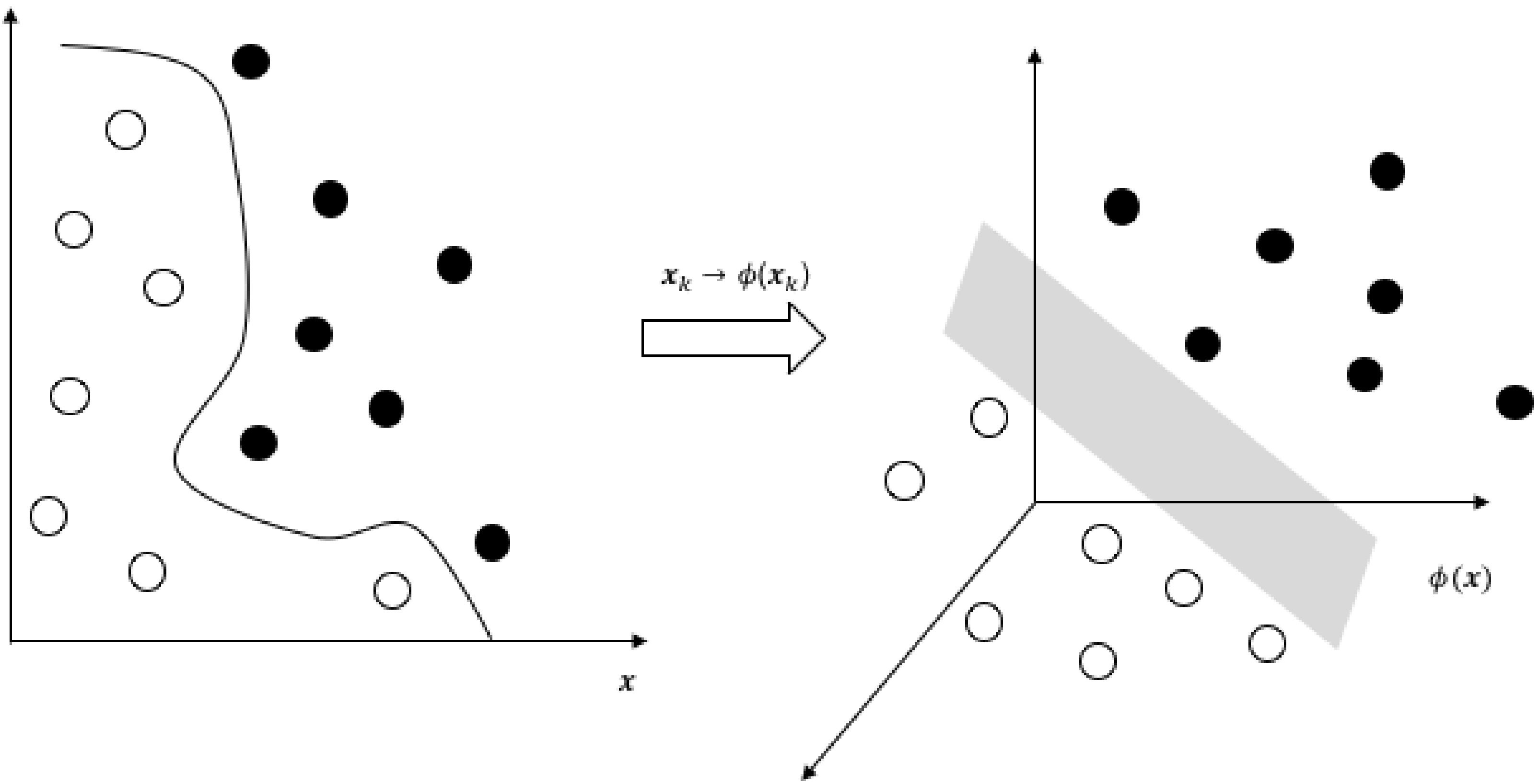
- Transforms x_i to a higher dimensional space using *kernel function* to consider non-linear case

MC-SVM: Modeling

- Kernel function
  - *Linear kernel function*
    - $K(x_i, x_j) = (x_j, x_i)$
  - *Polynomial kernel function with degree* d
    - $K(x_i, x_j) = (x_i^T x_j + 1)^d$
  - *Gaussian radial basis kernel function with* σ
    - $K(x_i, x_j) = \ exp\left(-\|x_i - x_j\|^2 / 2\sigma^2\right)$
  - *Sigmoid kernel function with* k *and* θ
    - $K(x_i, x_j) = tanh(k x_i^T x_j + \theta)$

MC-SVM: Modeling

- Non-linear SVM



MC-SVM: Modeling

- Solution (Using Sequential minimal optimization algorithm)

$$\alpha^* = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_i^*, \ldots, \alpha_n^*)^T$$

- SVM classifier function (i.e., decision function)

$$b^* = \sum_{i=1}^{n} \alpha_i^* y_i K(x_i, y_j)$$

$$f(x) = sgn(\sum_{i=1}^{n} \alpha_i^* y_i K(x_i, x) + b^*)$$

- MC-SVM can be solved by extending the binary-SVM model
  - One-versus-all (OVA)
  - One-versus-one (OVO)

Confusion matrix

N=165	Actual Positive(+)	Actual Negative(-)	
Predict Positive(+)	TP 100	FP 10	110
Predict Negative(-)	FN 5	TN 50	55
	105	60	

- True Positive (TP): Actual: pos. -> Predict: pos.
- True Negative (TN): Actual: neg. -> Predict: neg.
- False Positive (FP): Actual: neg. -> Predict: pos. (Type I error)
- False Negative (FN): Actual: pos. -> Predict: neg. (Type II error)

Confusion matrix

N=165	Actual Positive(+)	Actual Negative(-)	
Predict Positive(+)	TP 100	FP 10	110
Predict Negative(-)	FN 5	TN 50	55
	105	60	

- Precision: Predict: When it predicts pos. -> how often is it correct?
  - TP/(TP+FP) = 100/(110) = 0.91
- Recall: Actual: pos. -> how often does it predict pos.?
  - TP/(TP+FN) = 100/(100+5) =0.95 (Recall)
- Accuracy: How often is the classifier correct?
  - (TP+TN)/Total = (100+50)/165 =0.91

KDD Cup 1999 dataset: Features

Basic

Content

	데이터셋 속성	설명
1	Duration	연결 지속 시간
2	Protocol-type	프로토콜 종류 (예: TCP, UDP 등)
3	Service	서비스 종류 (예: HTTP, Telnet 등)
4	Flag	정상 또는 에러를 나타내는 플래그
5	Src_byte	출발지로부터의 데이터 크기
6	Dst_byte	목적지로부터의 데이터 크기
7	Land	출발지와 목적지의 주소가 같으면 1, 다르면 0
8	Wrong_fragment	프래그먼트 (fragment)오류의 개수
9	Urgen	Urgent 패킷의 개수
10	Count	두 호스트 간 2초 이상 연결을 지속한 접속 수
11	Srv_count	두 호스트 간 한 서비스로 2초 이상 연결을 지속한 접속 수
12	Serror_rate	SYN 에러율
13	Srv_serror_rate	서비스 SYN 에러율
14	Rerror_rate	REJ 에러율
15	Srv_rerror_rate	서비스 REJ 에러율
16	Same_srv_rate	접속중 같은 서비스 요청율
17	Diff_srv_rate	접속중 다른 서비스 요청율
18	Srv_diff_host_rate	다른 호스트 접속율
19	Dst_host_count	목적지 호스트 개수
20	Dst_host_srv_count	목적지 호스트 서비스 개수

KDD Cup 1999 dataset: Features

Content

21	Dst_host_same_srv_rate	목적지 호스트상 같은 서비스 비율
22	Dst_host_diff_srv_rate	목적지 호스트상 다른 서비스 비율
23	Dst_host_same_srv_port_rate	목적지 호스트상 같은 소스 포트 비율
24	Dst_host_diff_srv_host_rate	목적지 호스트상 다른 호스트율
25	Dst_host_serror_rate	목적지 호스트 SYN 에러율
26	Dst_host_rerror_rate	목적지 호스트 서비스 SYN 에러율
27	Dst_host_srv_rerror_rate	목적지 호스트 REJ 에러율
28	Dst_host_srv_rerror_rate	목적지 호스트 서비스 REJ 에러율

KDD Cup 1999 dataset: Mapping table

보안 위협 클래스	보안 위협 종류
Denial of Service (DoS)	Smurf, Land, Pod, Teardrop, Neptune, Back
Probing	Ipsweep, Nmap, Portsweep, Satan
User to Root (U2R)	Perl, Buffer_overflow, Rootkit, Loadmodule
Remote to Local (R2L)	Gess_pass, Imap, Multihop, Ftp_write, Phf, Spy, Warezmaster, Wareclient

	보안 위협 유형	플로우 수
1	Normal	78,010
2	Denial of Service (DoS)	3,712
3	Probing	3,796
4	User to Root (U2R)	35
5	Remote to Local (R2L)	1,125
합		86,678

KDD Cup 1999 dataset: Mapping table

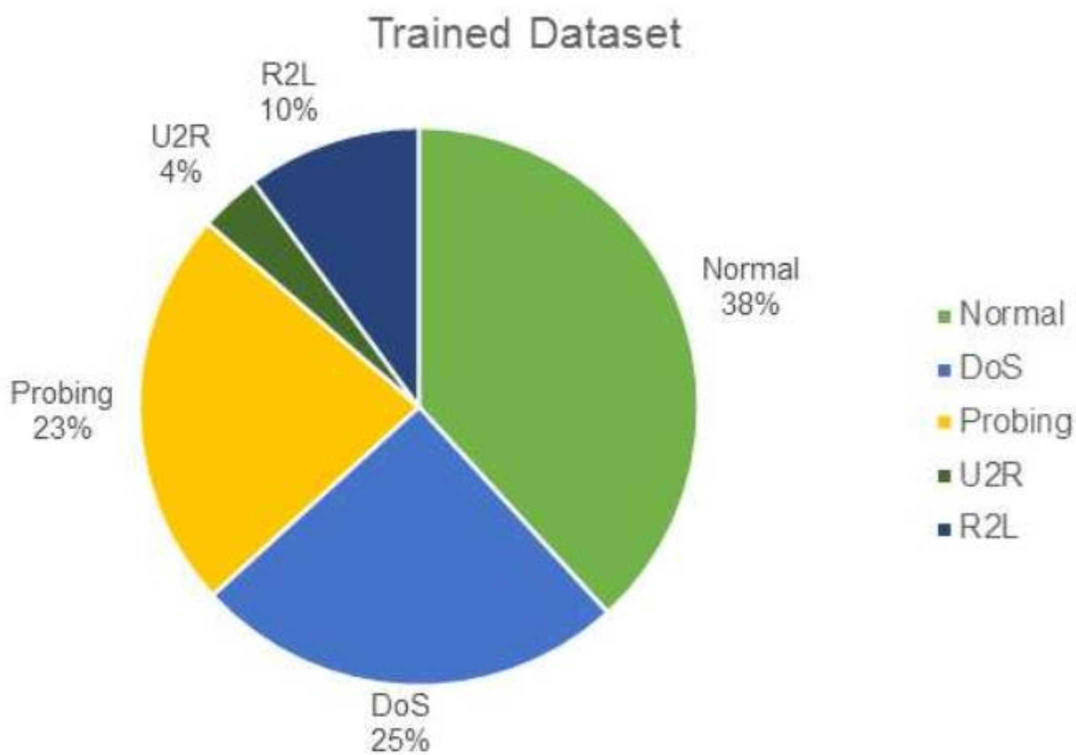
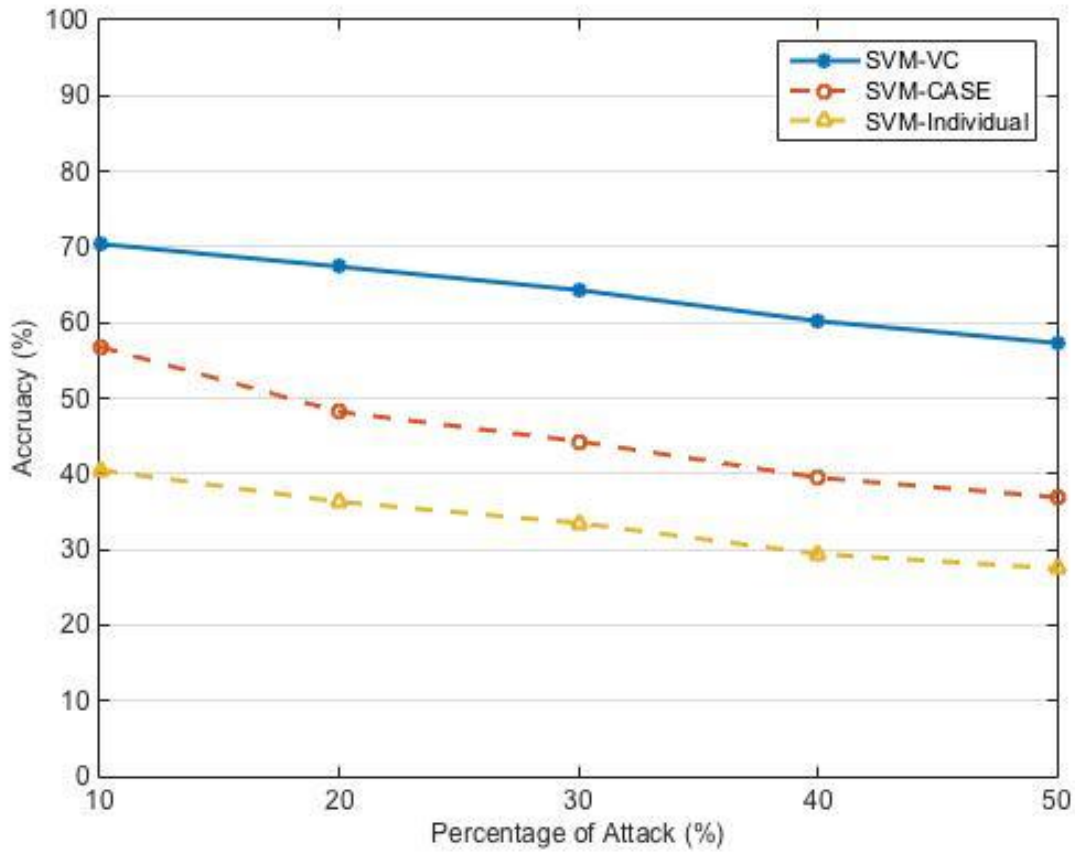


그림 12. 커널 함수 선택을 위한 학습 데이터셋 구성

Simulation results: Percentage of attack

- Accuracy



Simulation results: RoC

- RoC (Receiver Operating Characteristics)