Outside the Closed World: On Using Machine Learning for Network Intrusion Detection

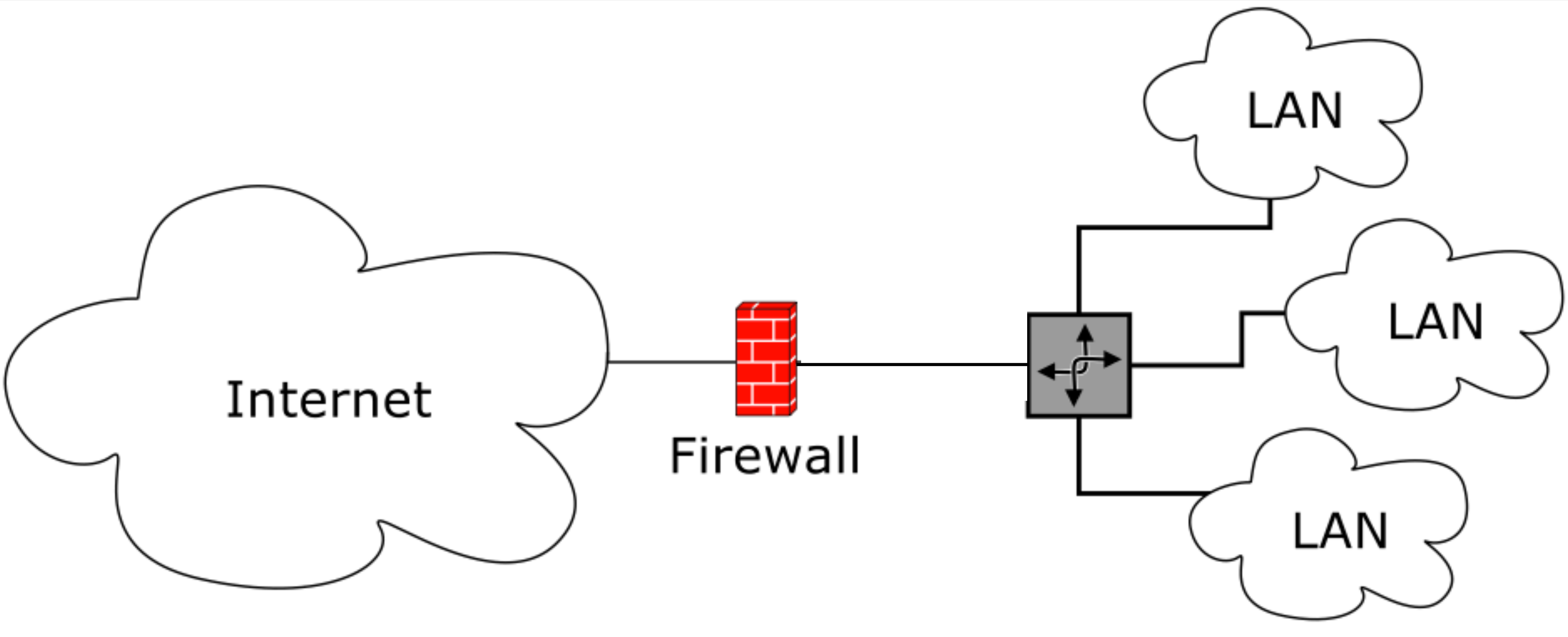
Robin Sommer
International Computer Science Institute, &
Lawrence Berkeley National Laboratory

Vern Paxson
International Computer Science Institute, &
University of California, Berkeley

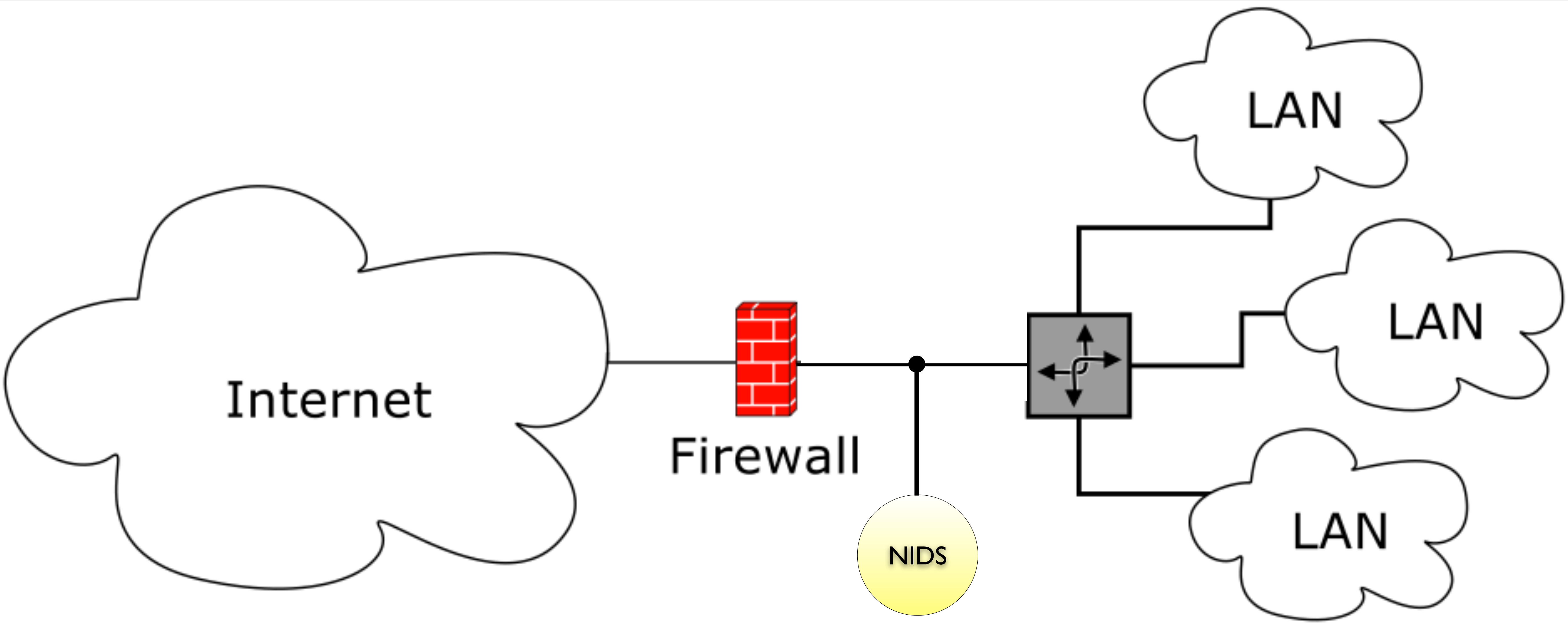
IEEE Symposium on Security and Privacy

May 2010

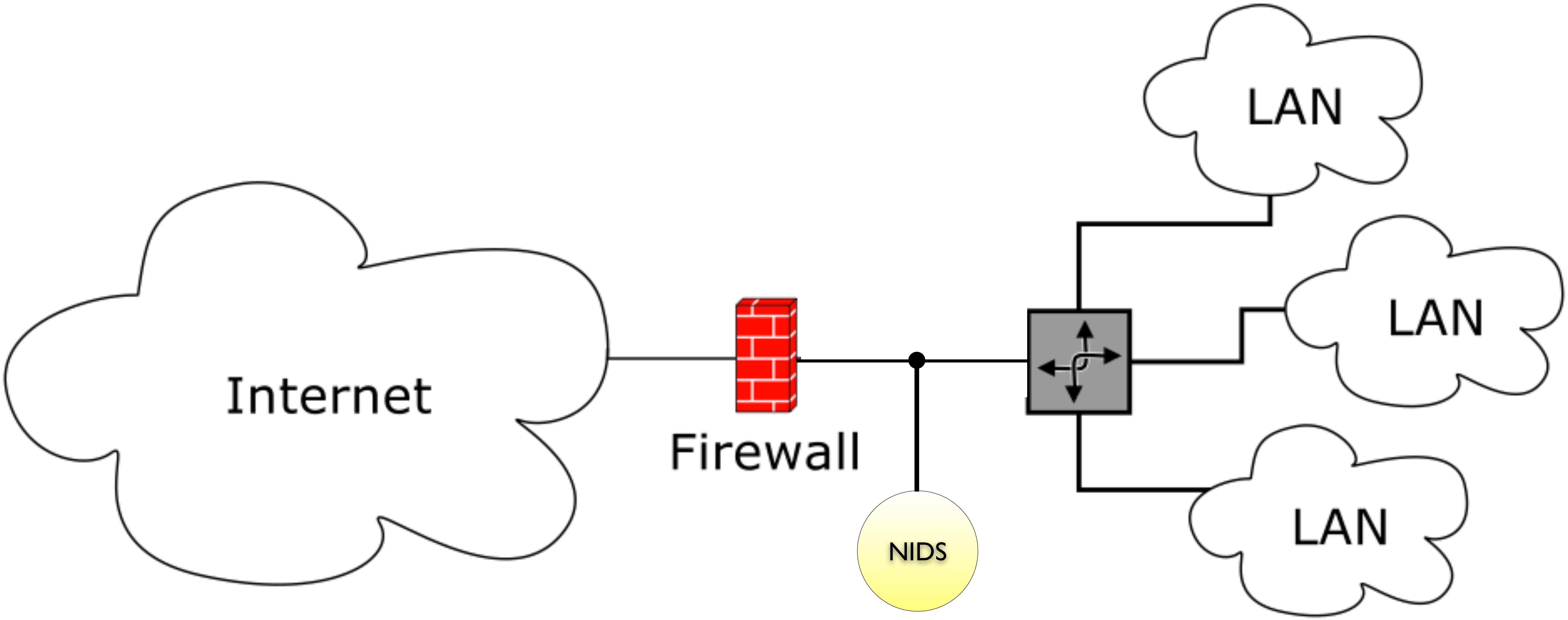
Network Intrusion Detection



Network Intrusion Detection

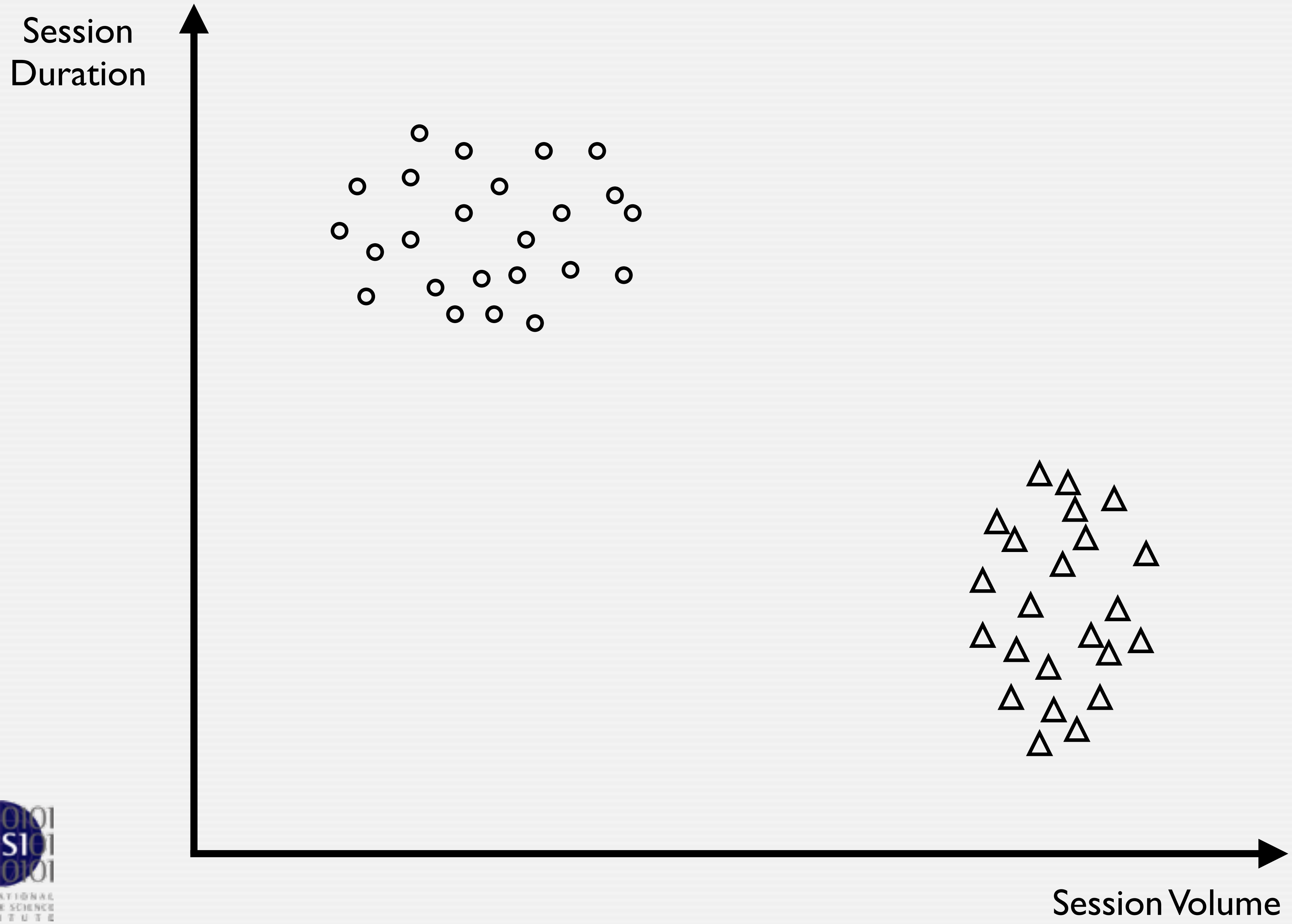


Network Intrusion Detection



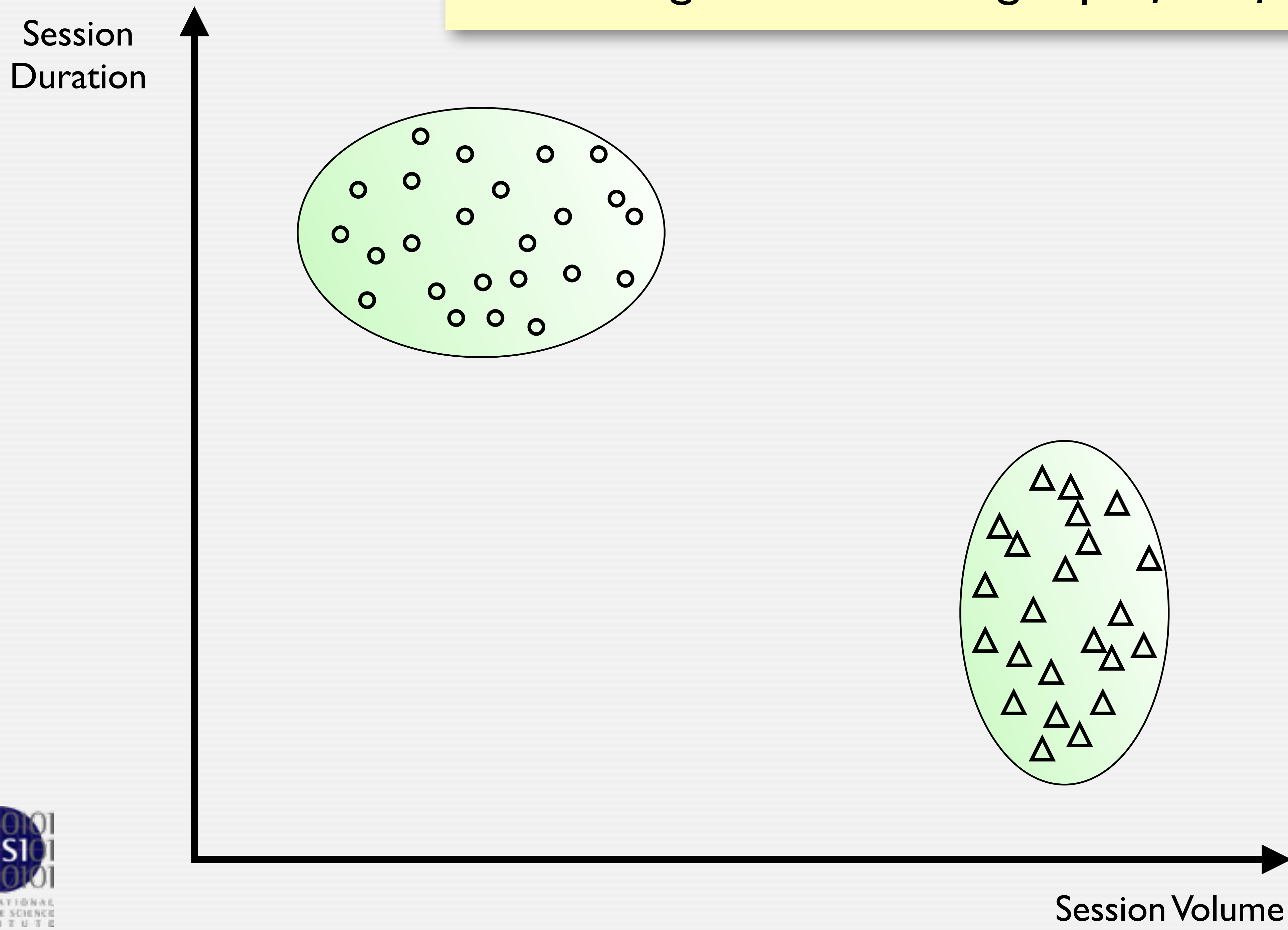
Detection Approaches: Misuse vs. Anomaly

Anomaly Detection



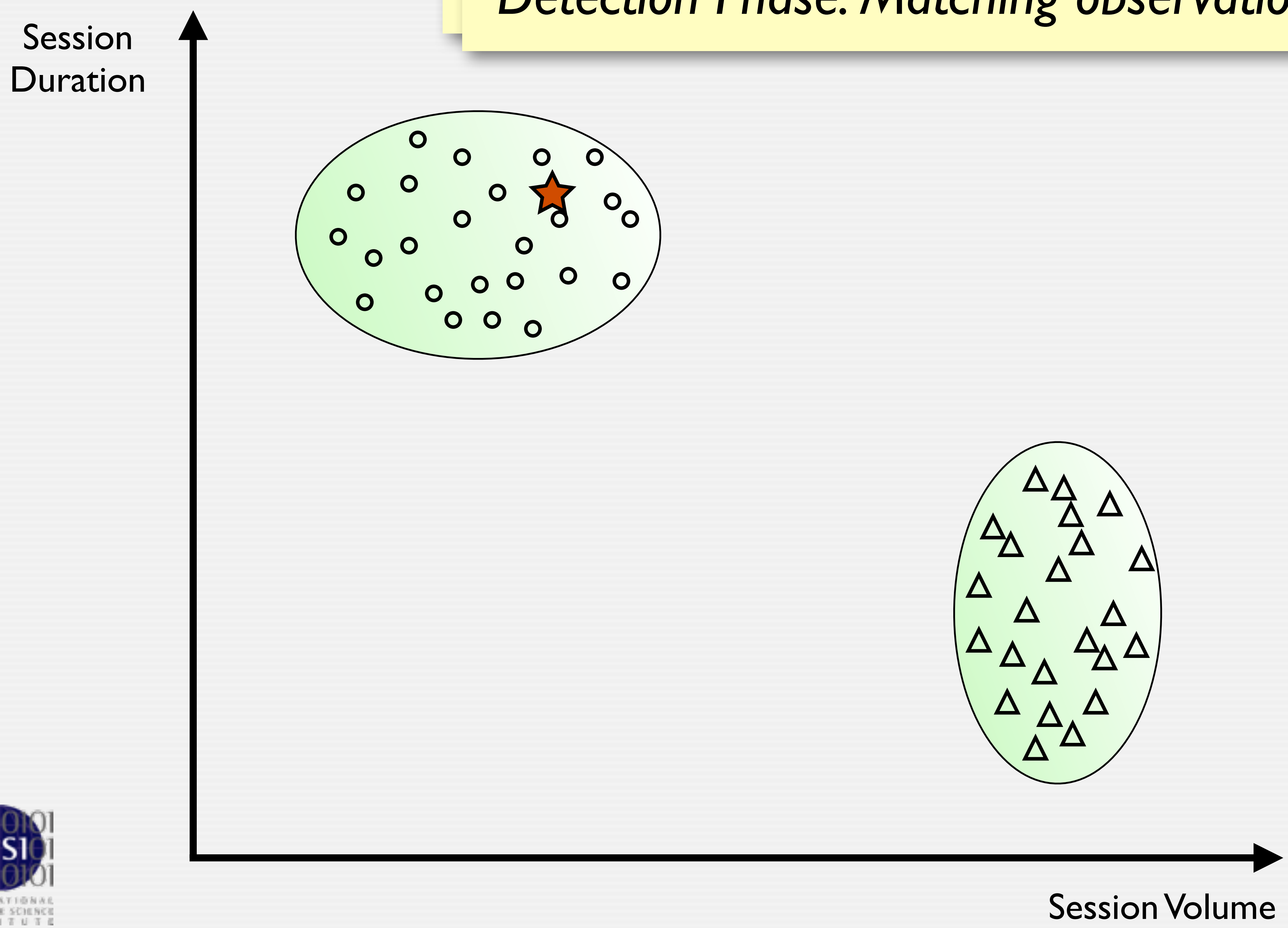
Anomaly Detection

Training Phase: Building a profile of normal activity.



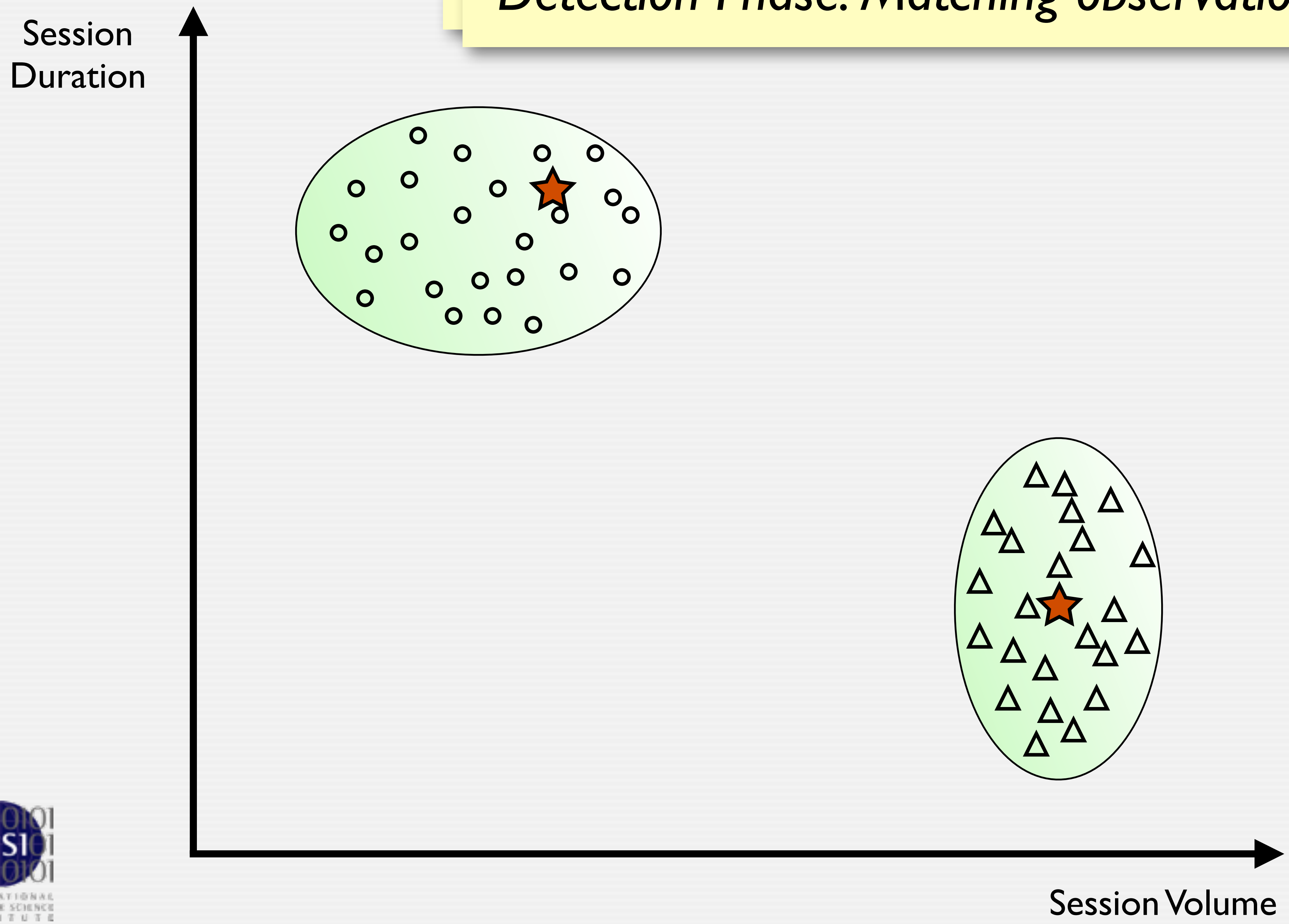
Anomaly Detection

Detection Phase: Matching observations against profile.



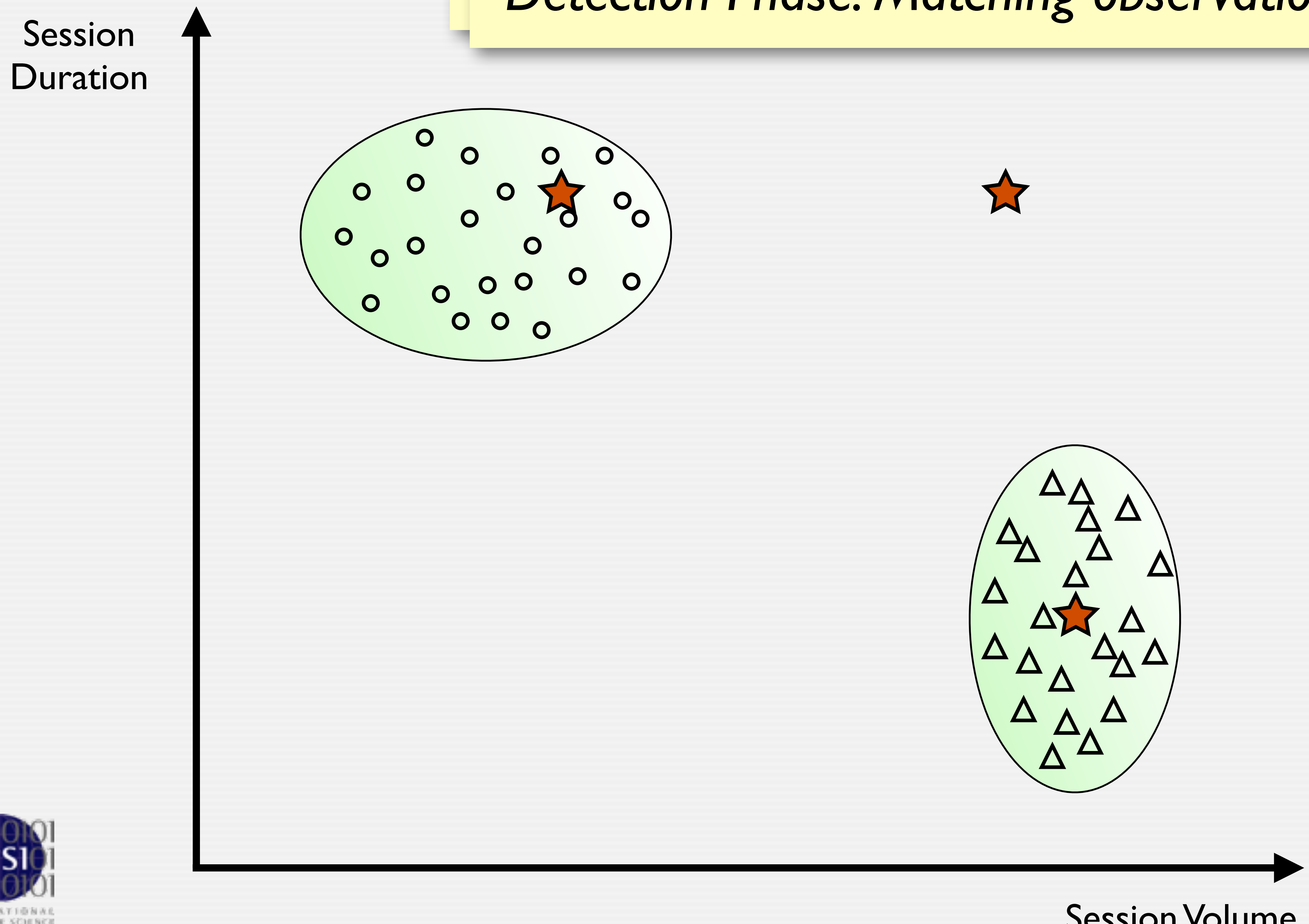
Anomaly Detection

Detection Phase: Matching observations against profile.



Anomaly Detection

Detection Phase: Matching observations against profile.



Anomaly Detection (2)

- Assumption: *Attacks exhibit characteristics that are different than those of normal traffic.*
- Originally introduced by Dorothy Denning in1987.
  - IDES: Host-level system building per-user profiles of activity.
  - Login frequency, password failures, session duration, resource consumption.

Anomaly Detection (2)

Technique Used	Section	References
Statistical Profiling using Histograms	Section 7.2.1	NIDES [Anderson et al. 1994; Anderson et al. 1995; Javitz and Valdes 1991], EMERALD [Porras and Neumann 1997], Yamanishi et al [2001; 2004], Ho et al. [1999], Kruegel at al [2002; 2003], Mahoney et al [2002; 2003; 2003; 2007], Sargor [1998]
Parametric Statistical Modeling	Section 7.1	Gwadera et al [2005b; 2004], Ye and Chen [2001]
Non-parametric Statistical Modeling	Section 7.2.2	Chow and Yeung [2002]
Bayesian Networks	Section 4.2	Siaterlis and Maglaris [2004], Sebyala et al. [2002], Valdes and Skinner [2000], Bronstein et al. [2001]
Neural Networks	Section 4.1	HIDE [Zhang et al. 2001], NSOM [Labib and Vemuri 2002], Smith et al. [2002], Hawkins et al. [2002], Kruegel et al. [2003], Manikopoulos and Papavassiliou [2002], Ramadas et al. [2003]
Support Vector Machines	Section 4.3	Eskin et al. [2002]
Rule-based Systems	Section 4.4	ADAM [Barbara et al. 2001a; Barbara et al. 2003; Barbara et al. 2001b], Fan et al. [2001], Helmer et al. [1998], Qin and Hwang [2004], Salvador and Chan [2003], Otey et al. [2003]
Clustering Based	Section 6	ADMIT [Sequeira and Zaki 2002], Eskin et al. [2002], Wu and Zhang [2003], Otey et al. [2003]
Nearest Neighbor based	Section 5	MINDS [Ertoz et al. 2004; Chandola et al. 2006], Eskin et al. [2002]
Spectral	Section 9	Shyu et al. [2003], Lakhina et al. [2005], Thottan and Ji [2003],Sun et al. [2007]
Information Theoretic	Section 8	Lee and Xiang [2001],Noble and Cook [2003]

Source: Chandola et al. 2009

Anomaly Detection (2)

Technique Used	Section	References
Statistical Profiling using Histograms	Section 7.2.1	NIDES [Anderson et al. 1994; Anderson et al. 1995; Javitz and Valdes 1991], EMERALD [Porras and Neumann 1997], Yamanishi et al [2001; 2004], Ho et al. [1999], Kruegel at al [2002; 2003], Mahoney et al [2002; 2003; 2003; 2007], Sargor [1998]
Parametric Statistical Modeling	Section 7.1	Gwadera et al [2005b; 2004], Ye and Chen [2001]
Non-parametric Statistical Modeling	Section 7.2.2	Chow and Yeung [2002]
Bayesian Networks	Section 4.2	Siaterlis and Maglaris [2004], Sebyala et al. [2002], Valdes and Skinner [2000], Bronstein et al. [2001]
Neural Networks	Section 4.1	HIDE [Zhang et al. 2001], NSOM [Labib and Vemuri 2002], Smith et al. [2002], Hawkins et al. [2002], Kruegel et al. [2003], Manikopoulos and Papavassiliou [2002], Ramadas et al. [2003]
Support Vector Machines	Section 4.3	Eskin et al. [2002]
Rule-based Systems	Section 4.4	ADAM [Barbara et al. 2001a; Barbara et al. 2003; Barbara et al. 2001b], Fan et al. [2001], Helmer et al. [1998], Qin and Hwang [2004], Salvador and Chan [2003], Otey et al. [2003]
Clustering Based	Section 6	ADMIT [Sequeira and Zaki 2002], Eskin et al. [2002], Wu and Zhang [2003], Otey et al. [2003]
Nearest Neighbor based	Section 5	MINDS [Ertoz et al. 2004; Chandola et al. 2006], Eskin et al. [2002]
Spectral	Section 9	Shyu et al. [2003], Lakhina et al. [2005], Thottan and Ji [2003],Sun et al. [2007]
Information Theoretic	Section 8	Lee and Xiang [2001],Noble and Cook [2003]

Source: Chandola et al. 2009

Features used

- packet sizes
- IP addresses
- ports
- header fields
- timestamps
- inter-arrival times
- session size
- session duration
- session volume
- payload frequencies
- payload tokens
- payload pattern
- ...

The Holy Grail ...

The Holy Grail ...

- Anomaly detection is extremely appealing.
  - Promises to find *novel* attacks without anticipating specifics.
  - It's *plausible*: machine learning works so well in other domains.

The Holy Grail ...

- Anomaly detection is extremely appealing.
  - Promises to find *novel* attacks without anticipating specifics.
  - It's *plausible*: machine learning works so well in other domains.
- But guess what's used *in operation*? Snort.
  - We find hardly any machine learning NIDS in real-world deployments.

The Holy Grail ...

- Anomaly detection is extremely appealing.
  - Promises to find *novel* attacks without anticipating specifics.
  - It's *plausible*: machine learning works so well in other domains.
- But guess what's used *in operation*? Snort.
  - We find hardly any machine learning NIDS in real-world deployments.
- Could using machine learning be harder than it appears?

Why is Anomaly Detection Hard?

The intrusion detection domain faces challenges that make it fundamentally different from other fields.

Why is Anomaly Detection Hard?

The intrusion detection domain faces challenges that make it fundamentally different from other fields.

Outlier detection and the high costs of errors

How do we find the opposite of normal?

Interpretation of results

What does that anomaly *mean*?

Evaluation

How do we make sure it actually works?

Training data

What do we train our system with?

Evasion risk

Can the attacker mislead our system?

Why is Anomaly Detection Hard?

The intrusion detection domain faces challenges that make it fundamentally different from other fields.

Outlier detection and the high costs of errors

How do we find the opposite of normal?

Interpretation of results

What does that anomaly *mean*?

Evaluation

How do we make sure it actually works?

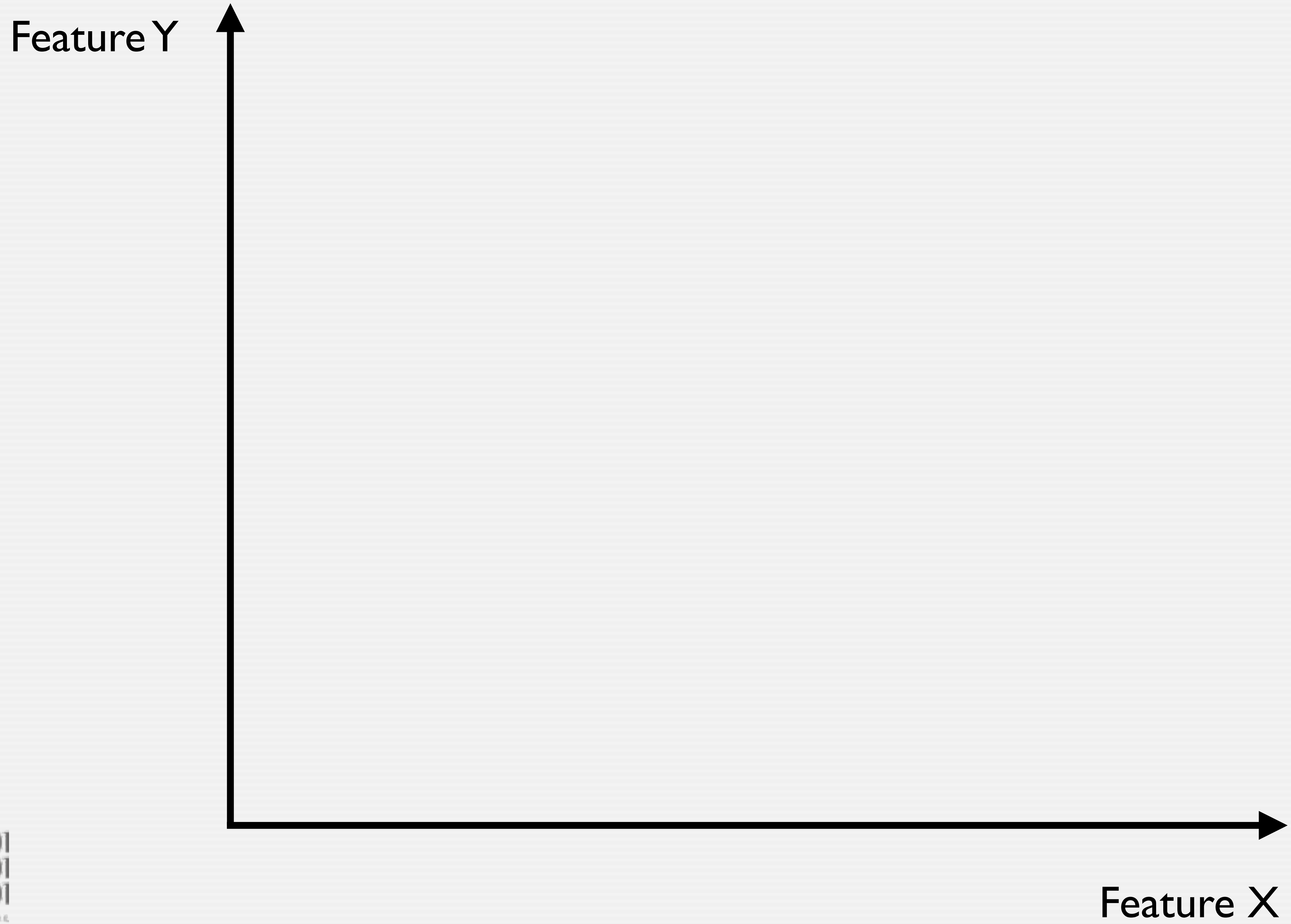
Training data

What do we train our system with?

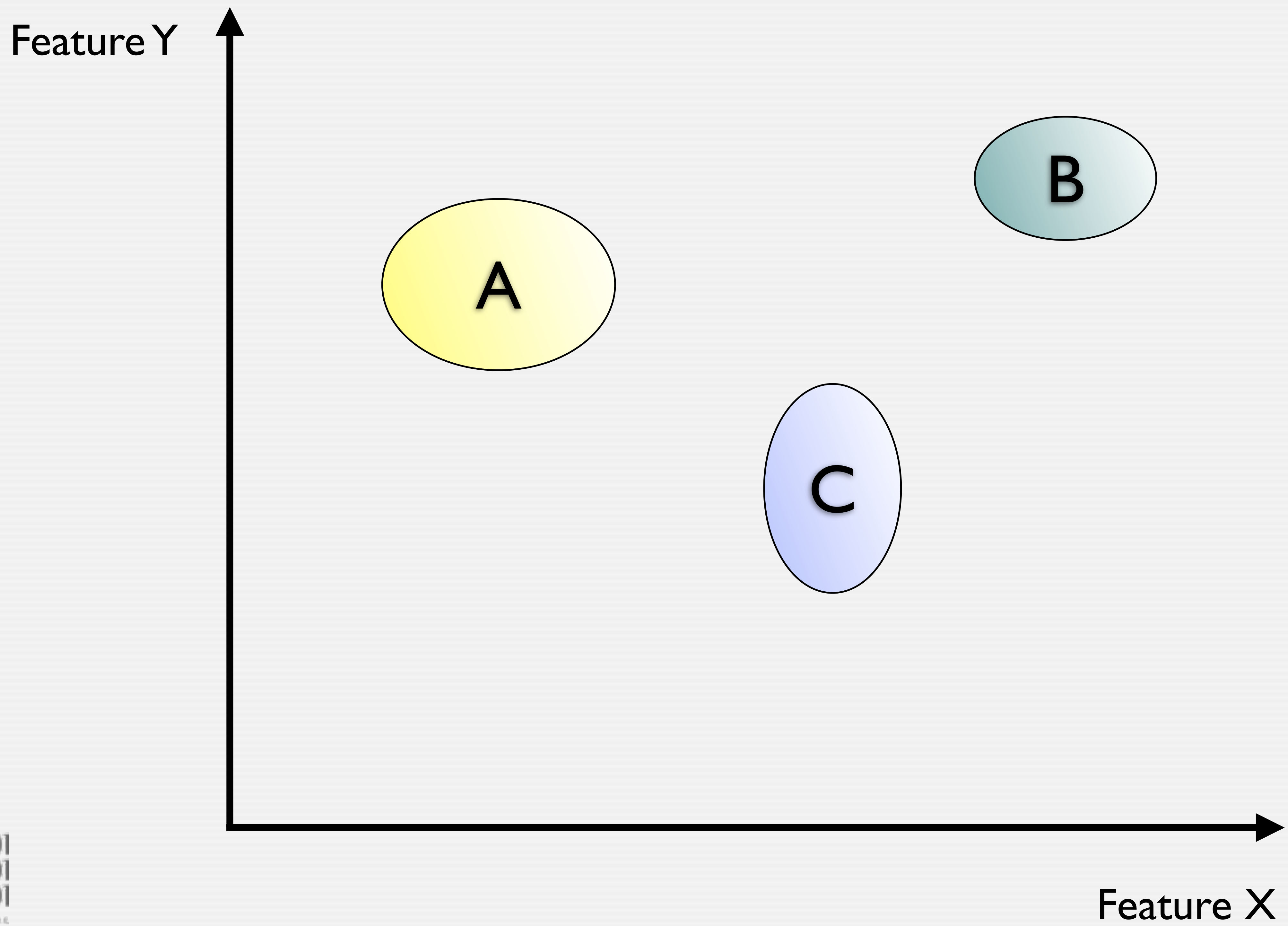
Evasion risk

Can the attacker mislead our system?

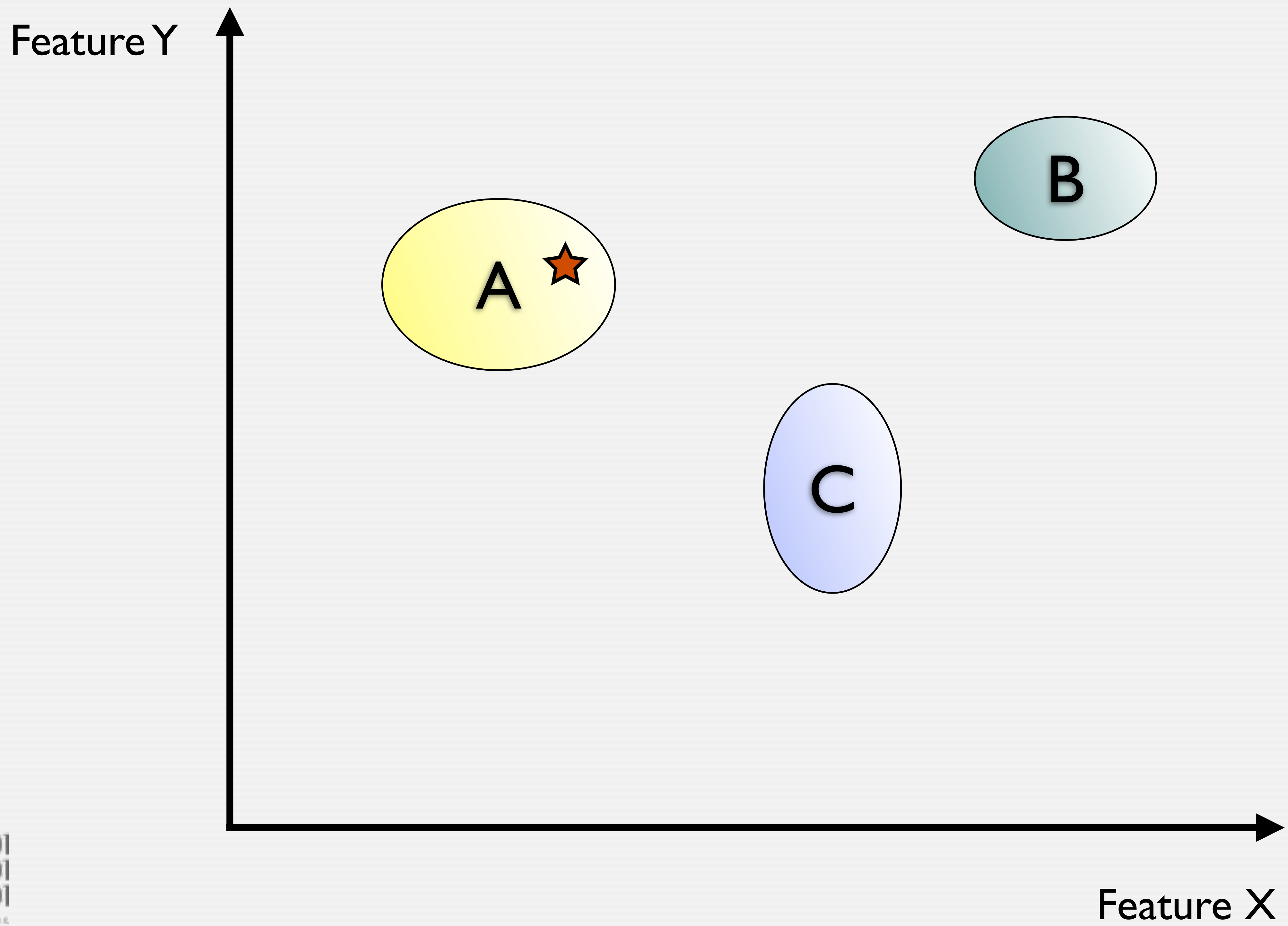
Machine Learning for Classification



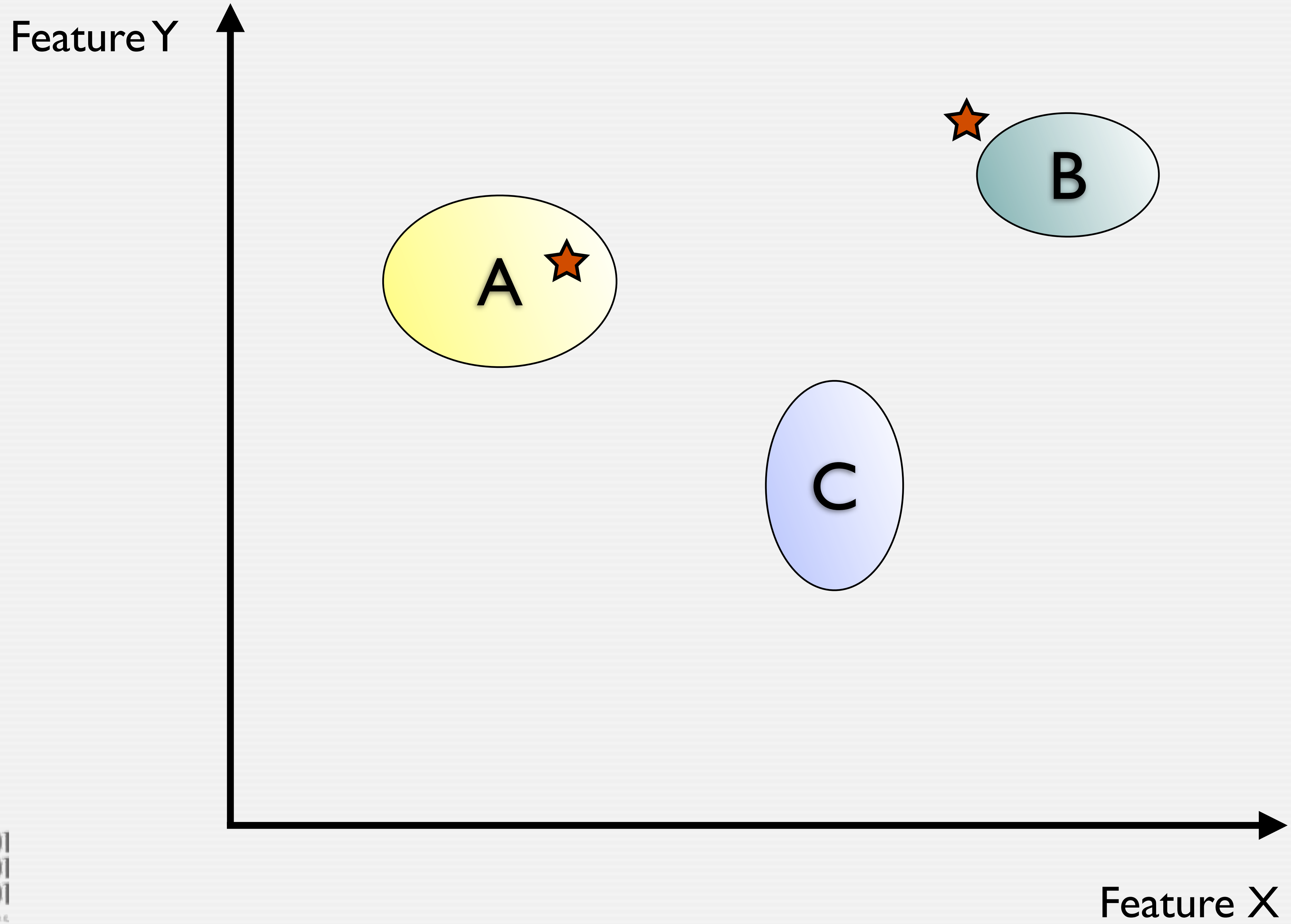
Machine Learning for Classification



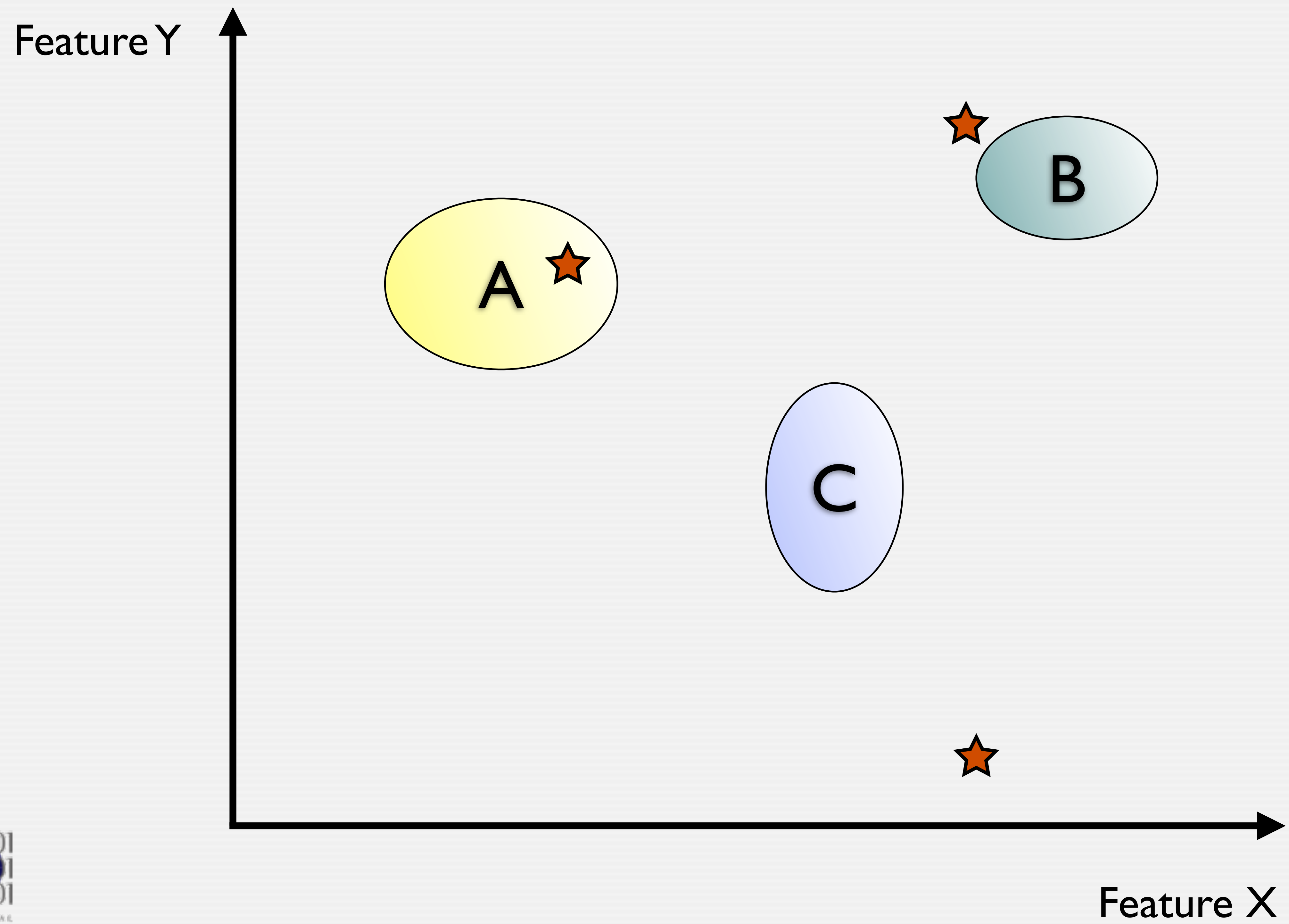
Machine Learning for Classification



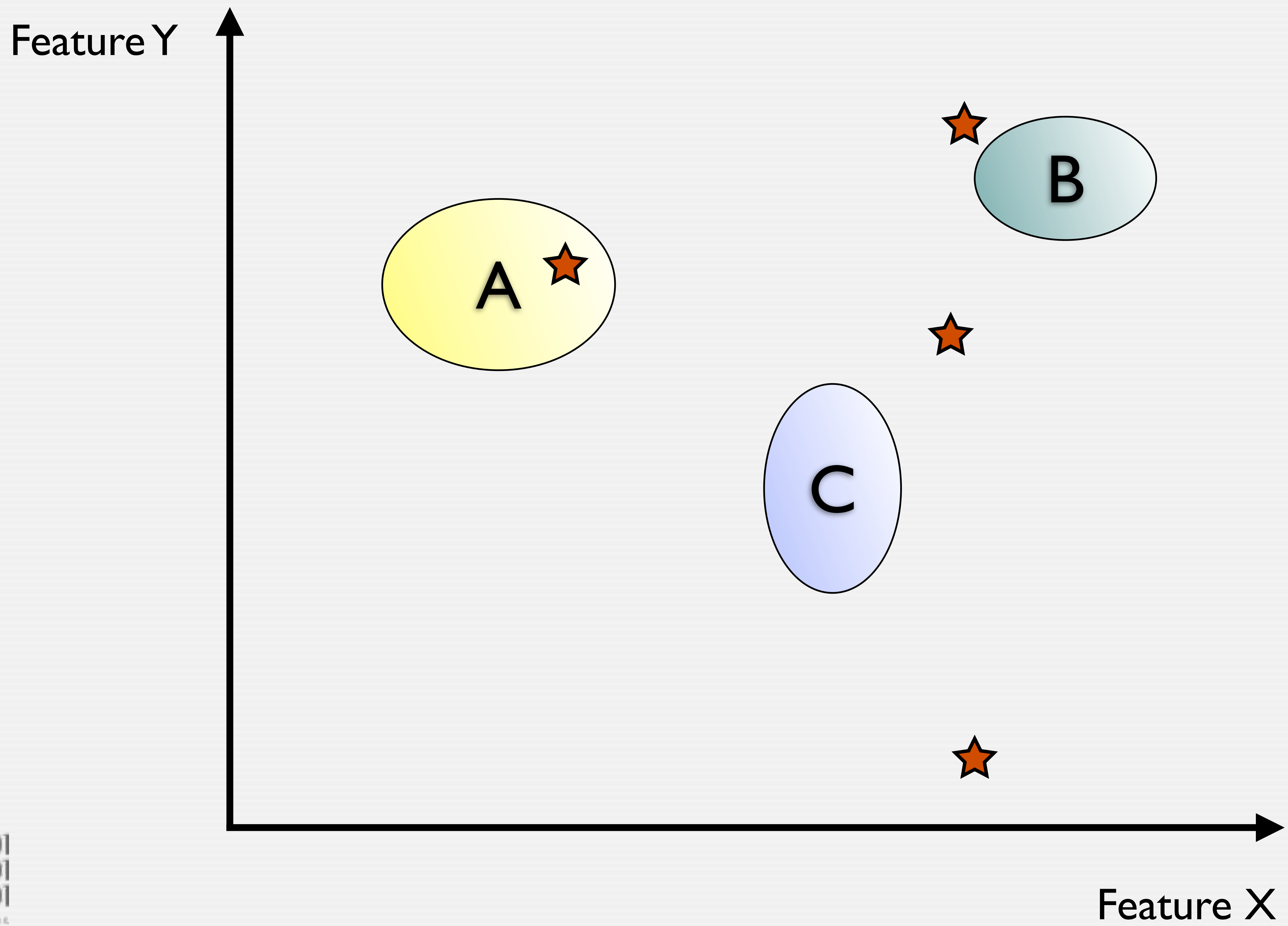
Machine Learning for Classification



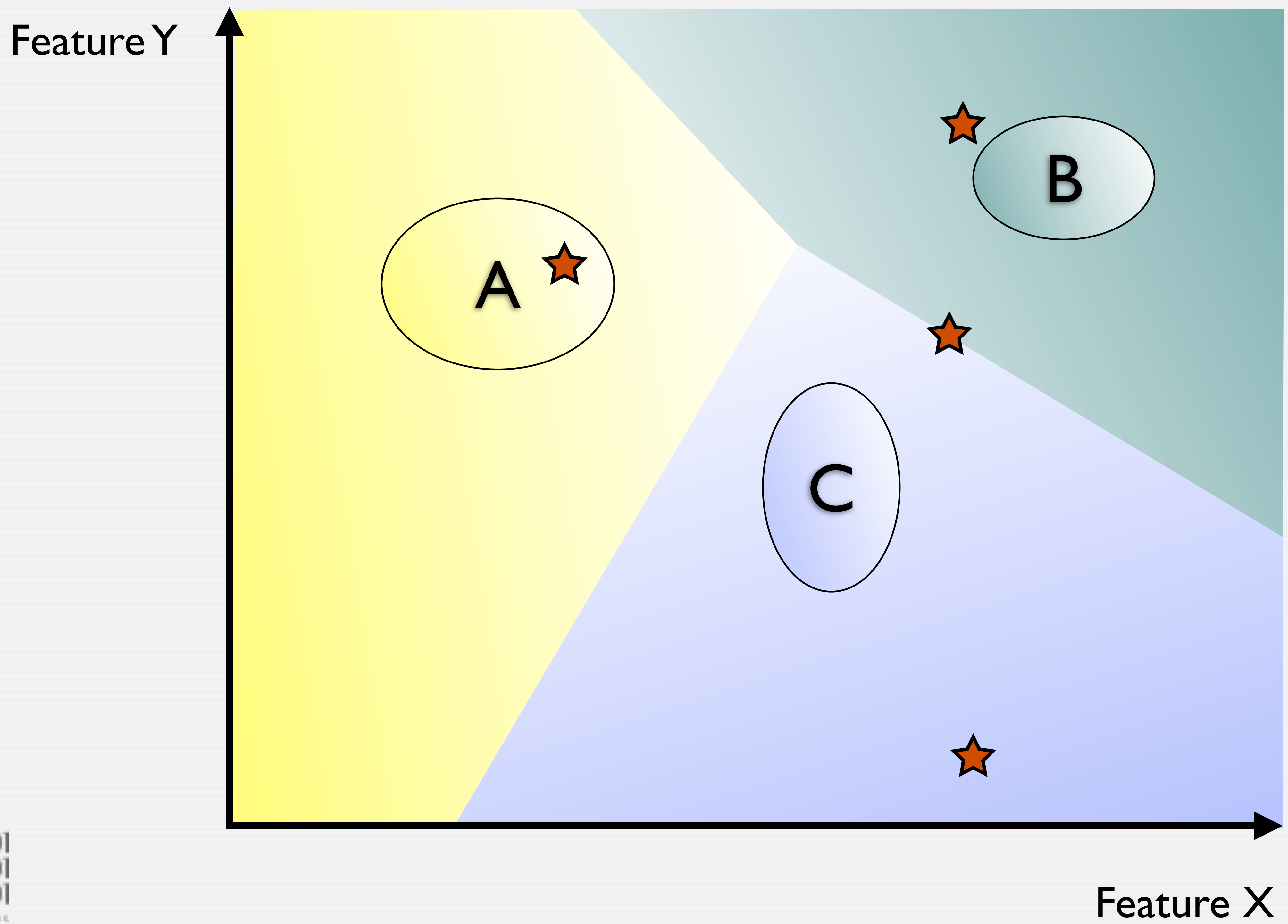
Machine Learning for Classification



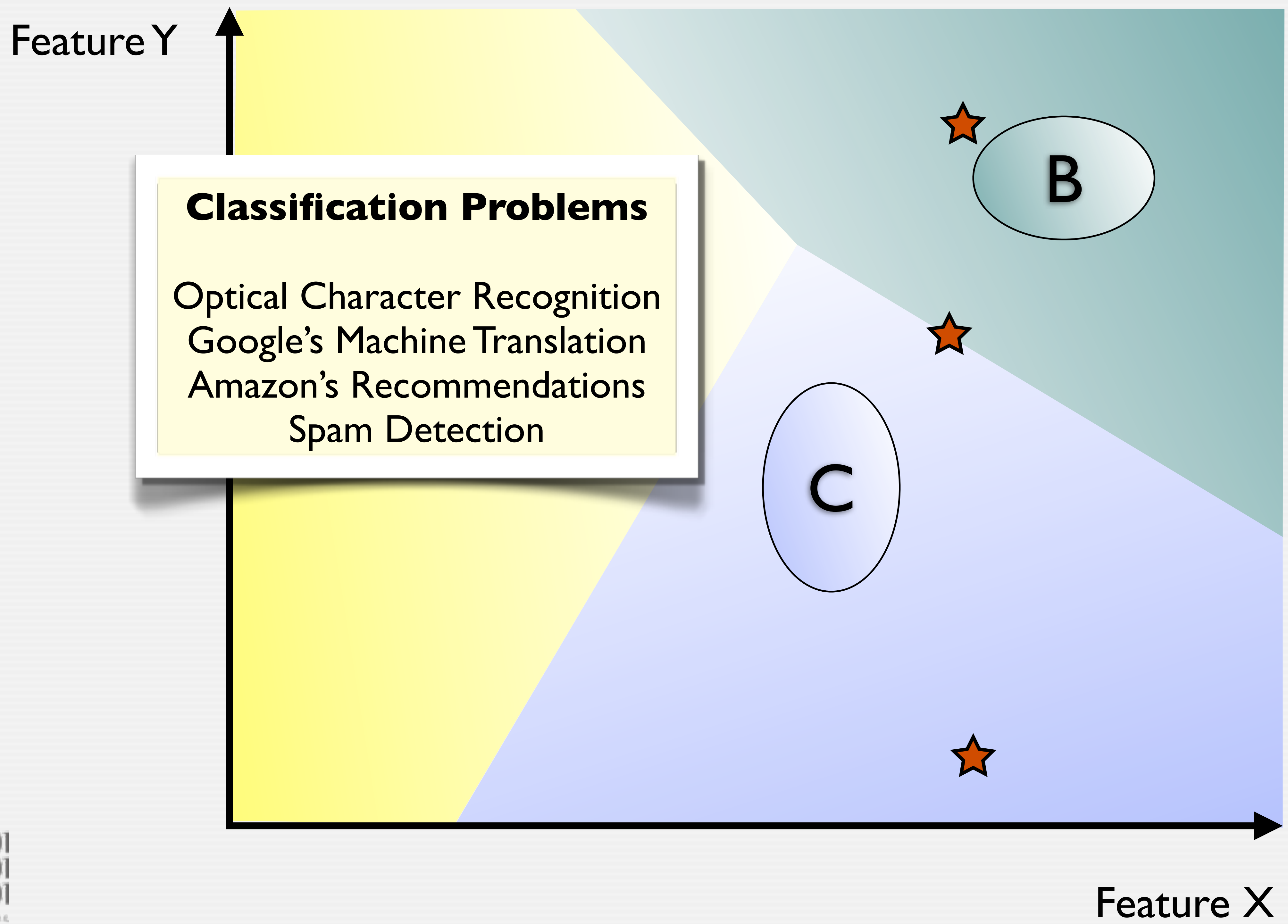
Machine Learning for Classification



Machine Learning for Classification



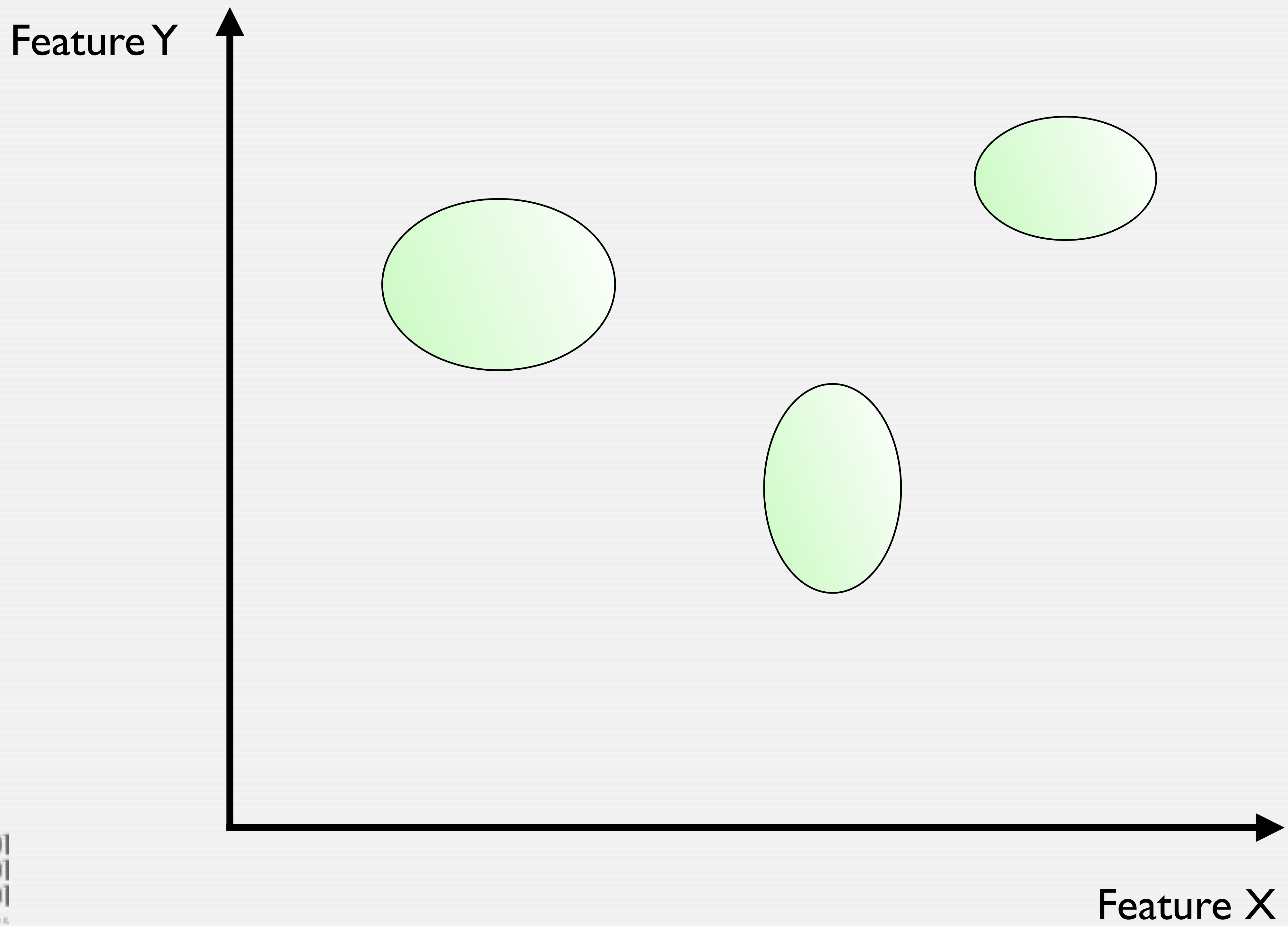
Machine Learning for Classification



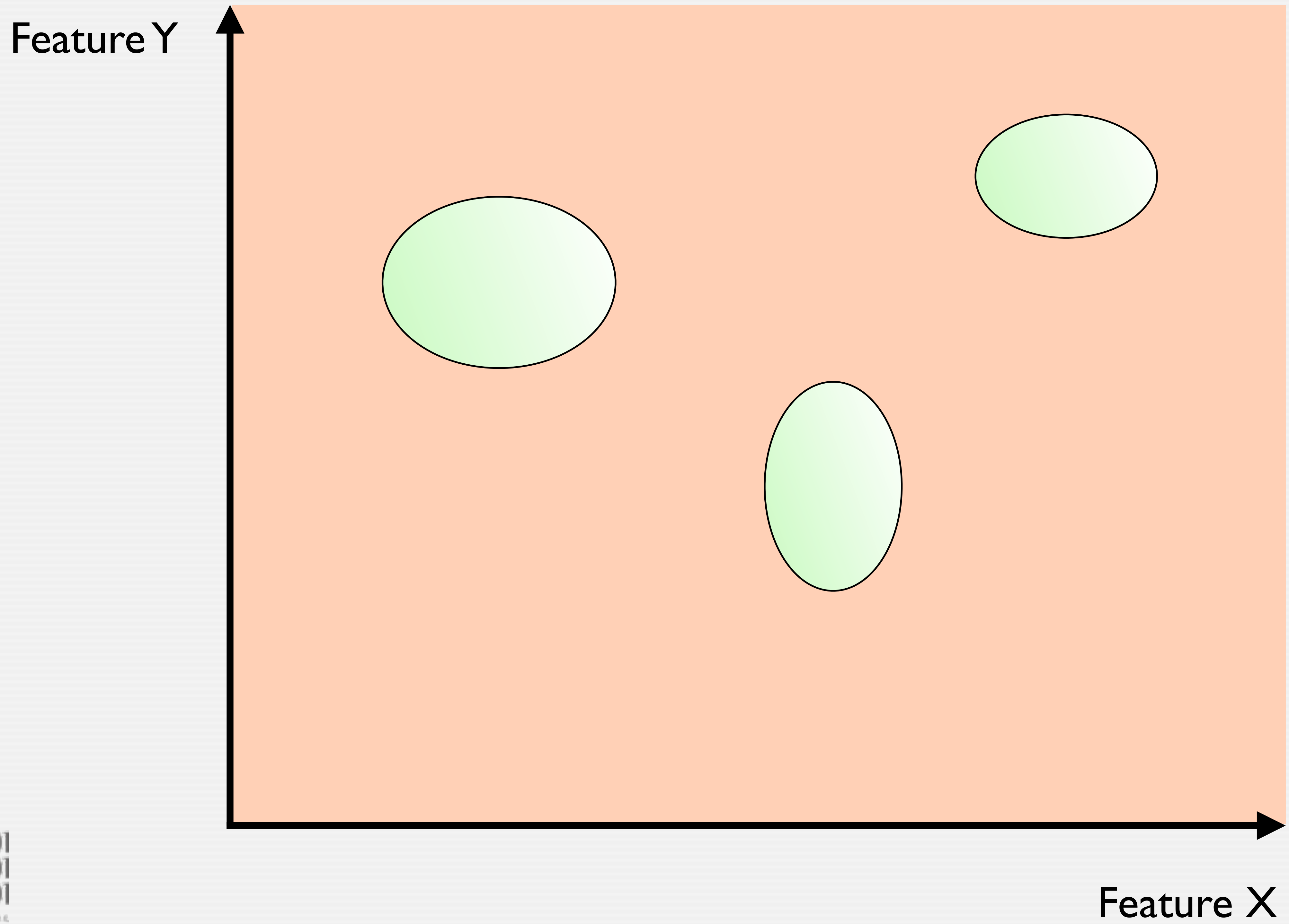
Classification Problems

Optical Character Recognition
Google's Machine Translation
Amazon's Recommendations
Spam Detection

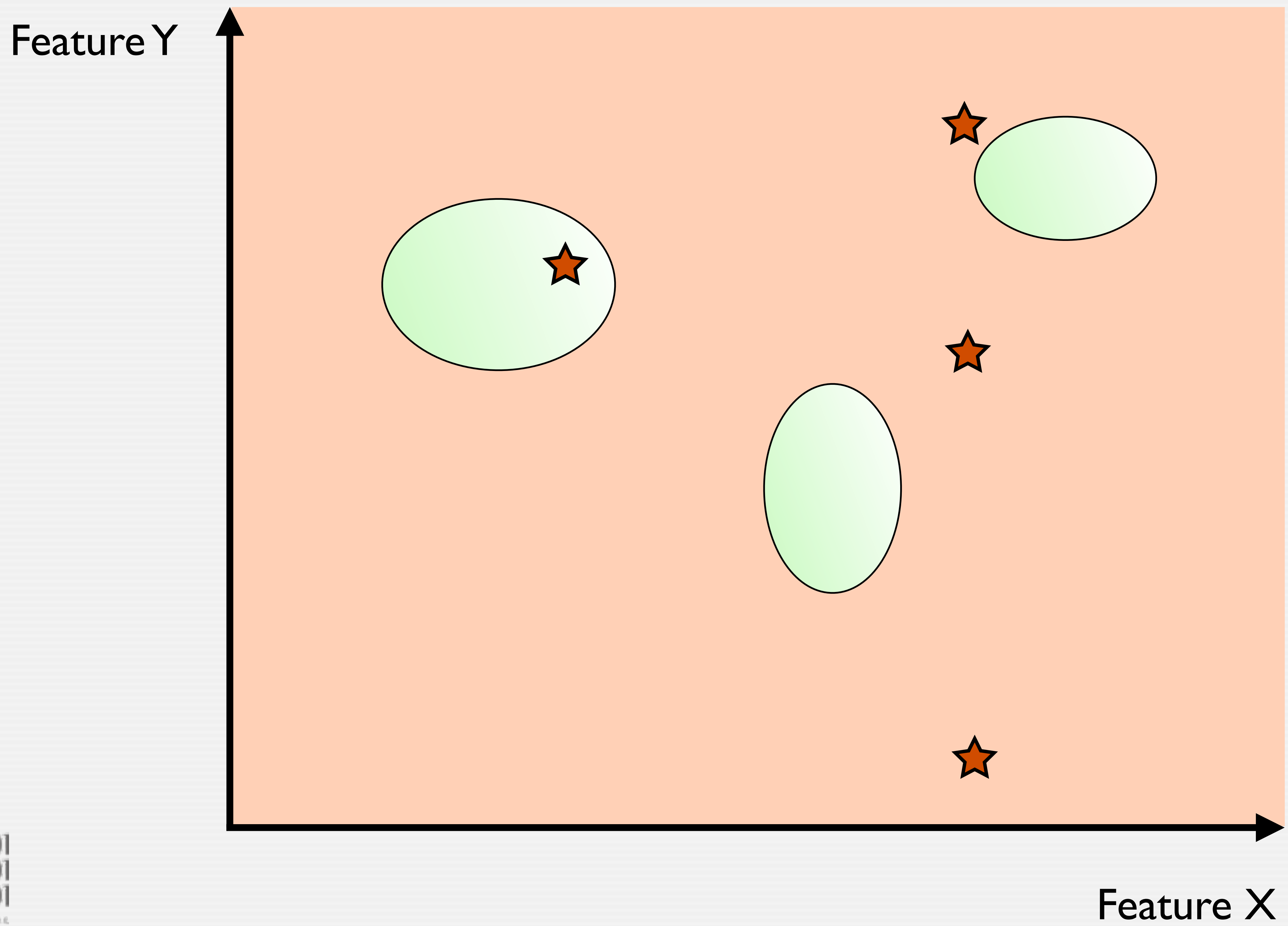
Outlier Detection



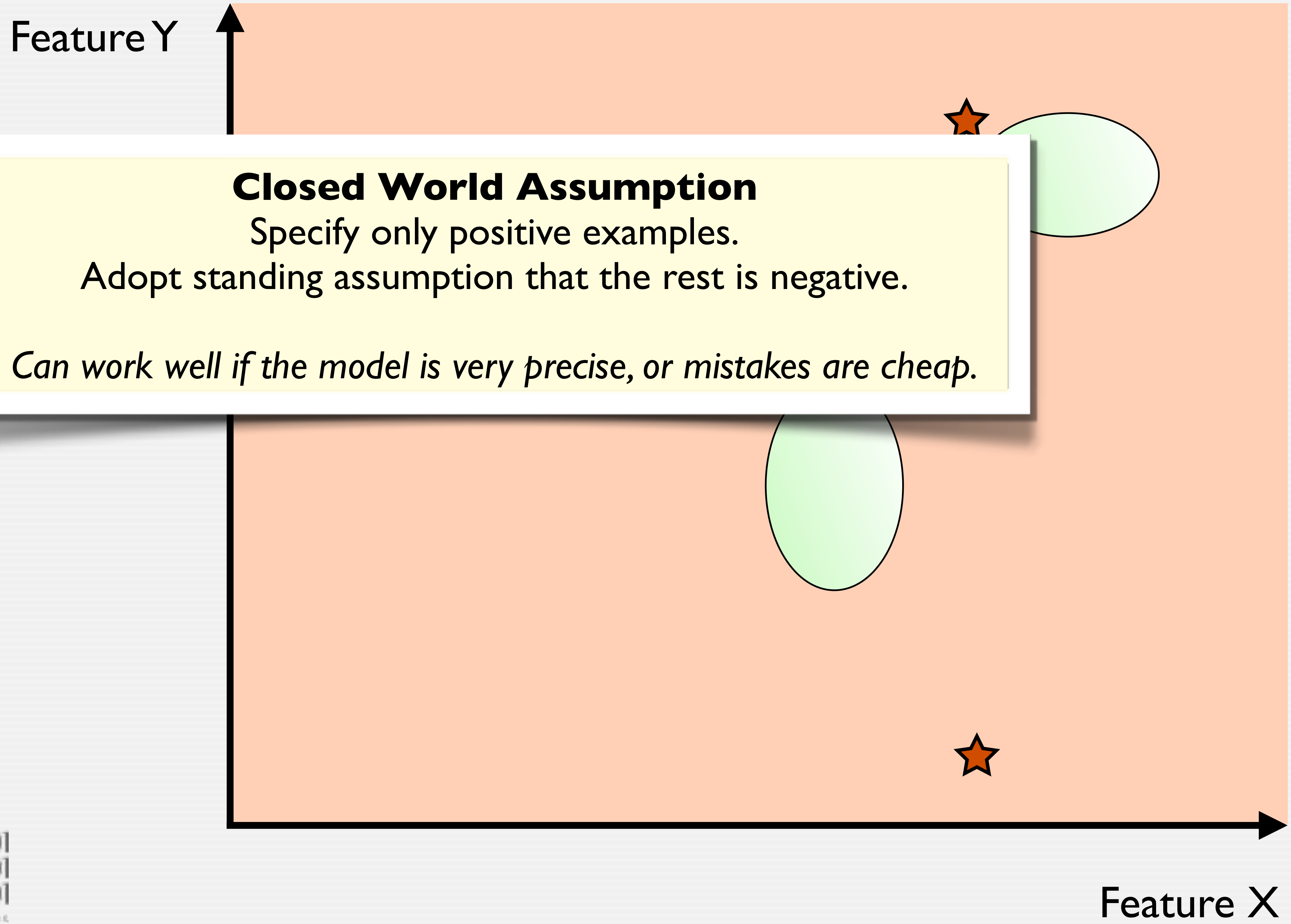
Outlier Detection



Outlier Detection



Outlier Detection



Closed World Assumption

Specify only positive examples.

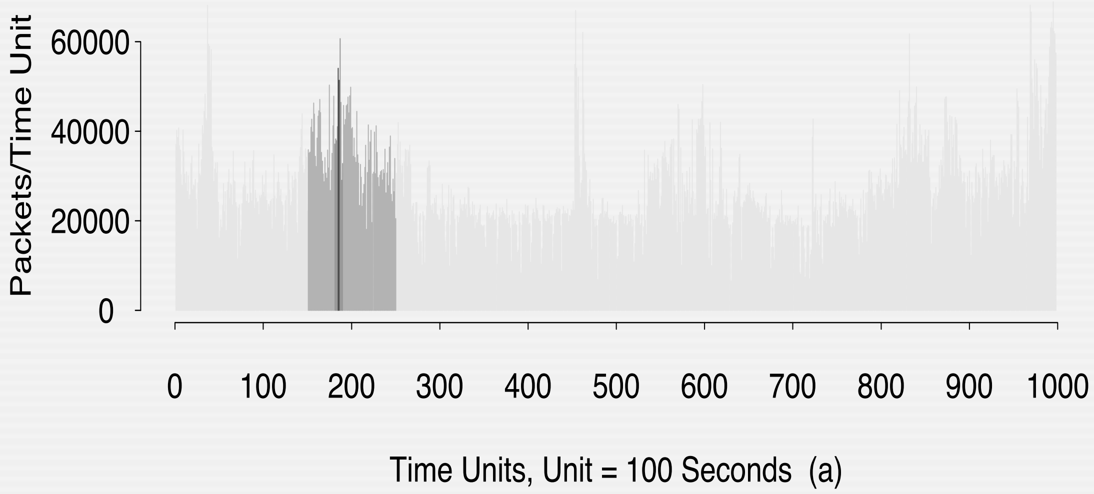
Adopt standing assumption that the rest is negative.

Can work well if the model is very precise, or mistakes are cheap.

What *is* Normal?

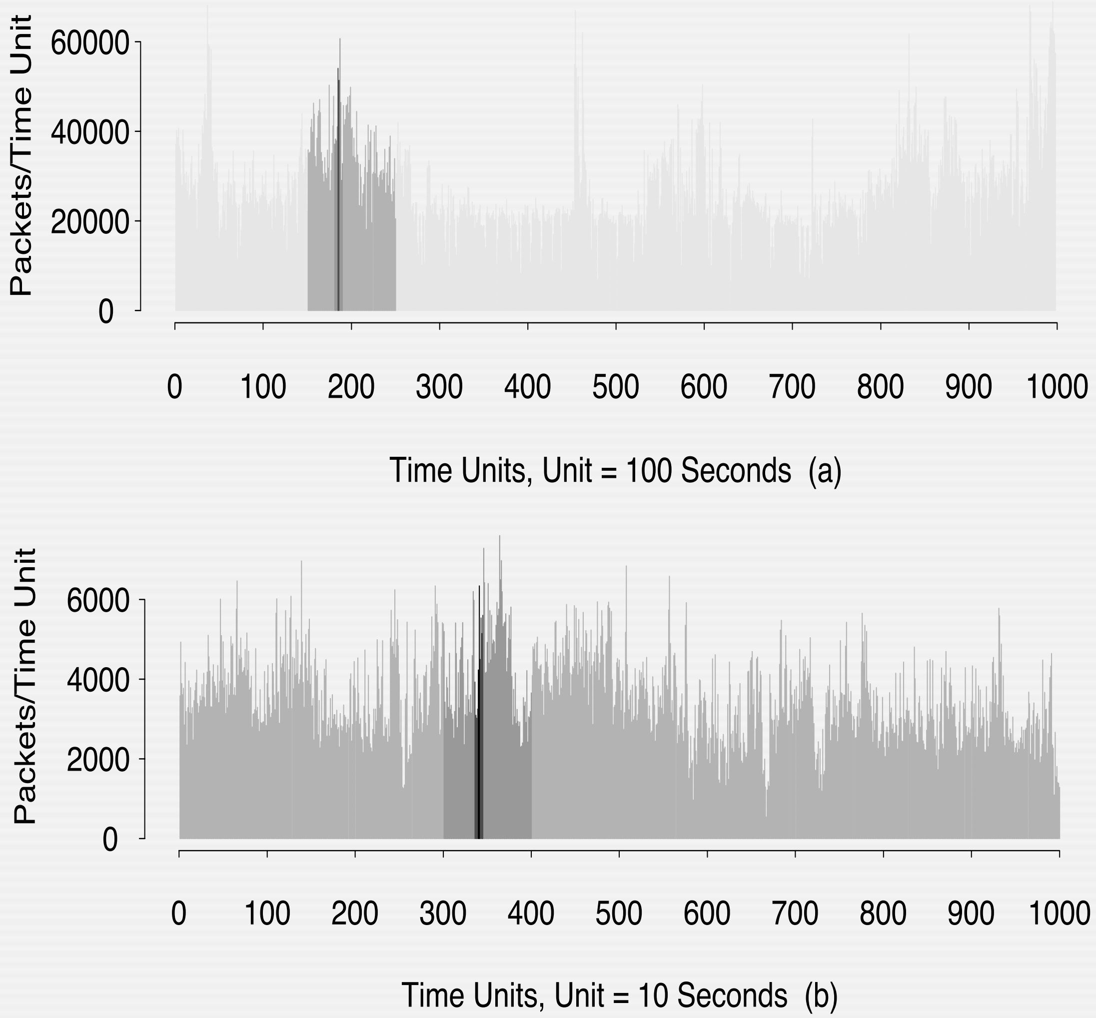
- Finding a stable notion of normal is hard for networks.
- Network traffic is composed of *many* individual sessions.
  - Leads to enormous variety and unpredictable behavior.
  - Observable on all layers of the protocol stack.

Self-Similarity of Ethernet Traffic



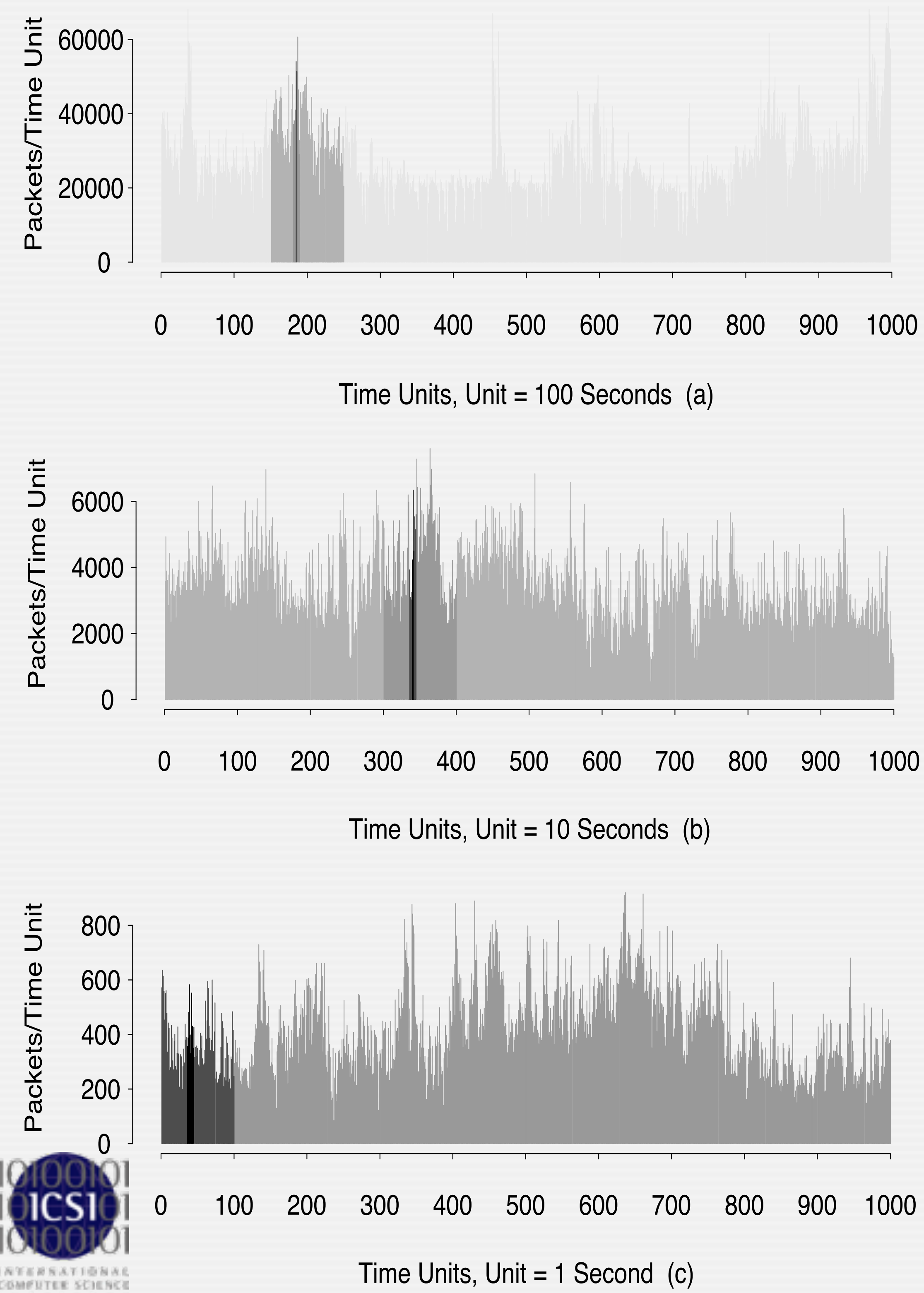
Source: LeLand et al. 1995

Self-Similarity of Ethernet Traffic



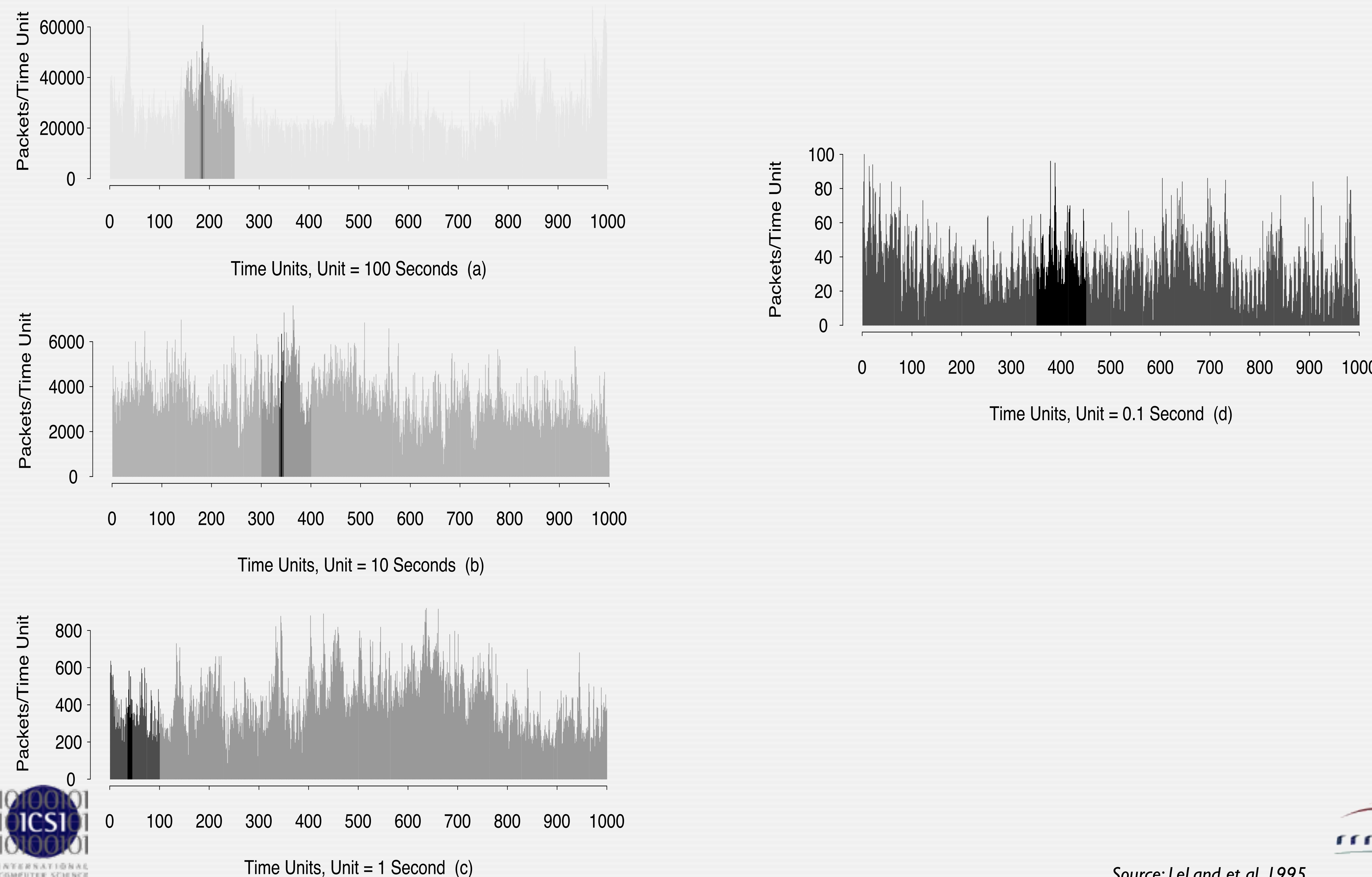
Source: LeLand et al. 1995

Self-Similarity of Ethernet Traffic



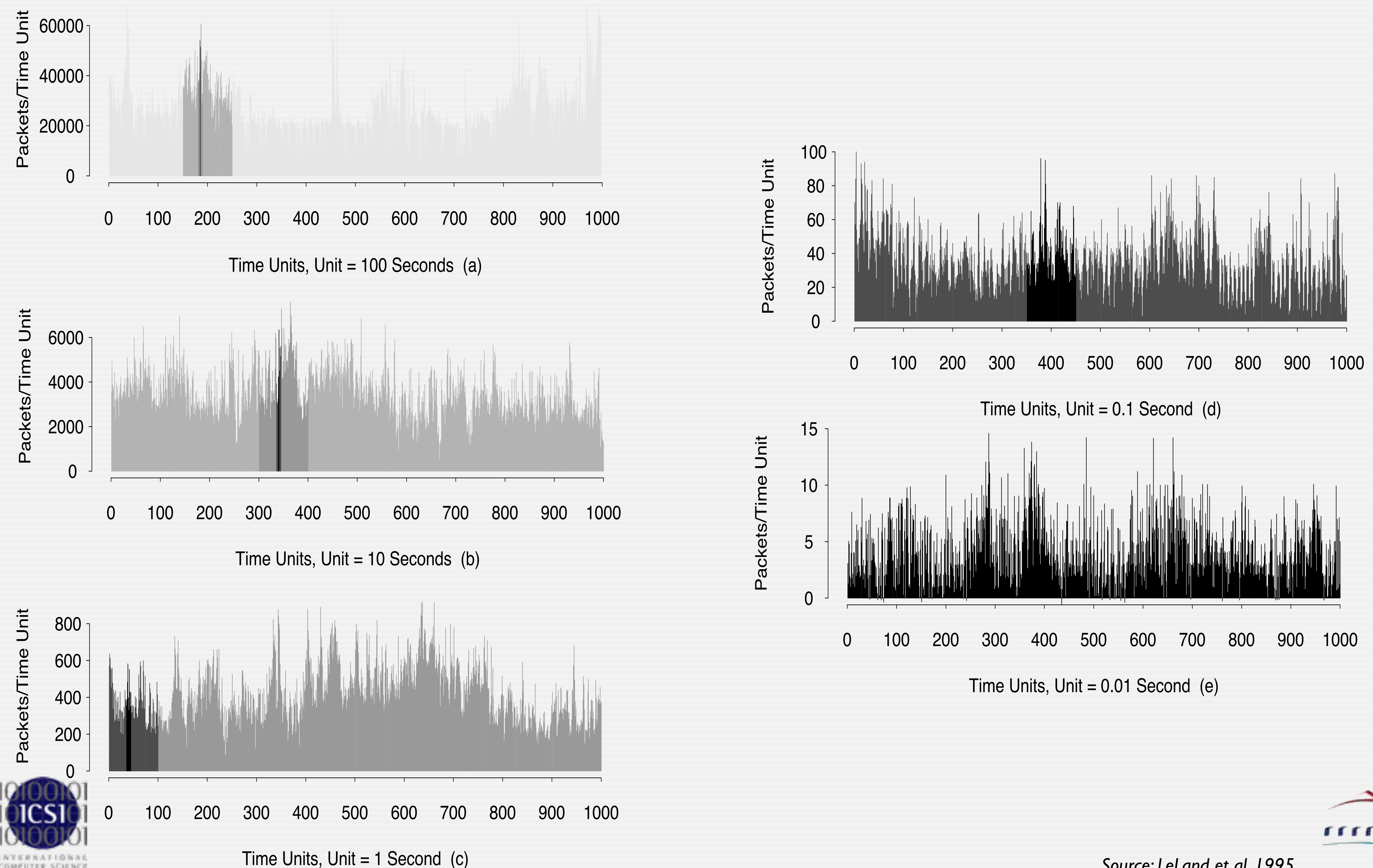
Source: LeLand et al. 1995

Self-Similarity of Ethernet Traffic



Source: LeLand et al. 1995

Self-Similarity of Ethernet Traffic



Source: LeLand et al. 1995

One Day of Crud at ICSI

Postel's Law: *Be strict in what you send and liberal in what you accept ...*

One Day of Crud at ICSI

Postel's Law: *Be strict in what you send and liberal in what you accept ...*

active-connection-reuse	DNS-label-len-gt-pkt	HTTP-chunked-multipart	possible-split-routing
bad-Ident-reply	DNS-label-too-long	HTTP-version-mismatch	SYN-after-close
bad-RPC	DNS-RR-length-mismatch	illegal-%-at-end-of-URI	SYN-after-reset
bad-SYN-ack	DNS-RR-unknown-type	inappropriate-FIN	SYN-inside-connection
bad-TCP-header-len	DNS-truncated-answer	IRC-invalid-line	SYN-seq-jump
base64-illegal-encoding	DNS-len-lt-hdr-len	line-terminated-with-single-CR	truncated-NTP
connection-originator-SYN-ack	DNS-truncated-RR-rdlength	malformed-SSH-identification	unescaped-%-in-URI
data-after-reset	double-%-in-URI	no-login-prompt	unescaped-special-URI-char
data-before-established	excess-RPC	NUL-in-line	unmatched-HTTP-reply
too-many-DNS-queries	FIN-advanced-last-seq	POP3-server-sending-client-commands	window-recision
DNS-label-forward-compress-	fragment-with-DF		***155K in total!***

ICSI INTERNATIONAL COMPUTER SCIENCE INSTITUTE

BERKELEY LAB

What *is* Normal?

- Finding a stable notion of normal is hard for networks.
- Network traffic is composed of *many* individual sessions.
  - Leads to enormous variety and unpredictable behavior.
  - Observable on all layers of the protocol stack.

- Violates an implicit assumption: Outliers are attacks!
- Ignoring this leads to a *semantic gap*
  - Disconnect between what the system reports and what the operator wants.
  - Root cause for the common complaint of "too many false positives".
- Each mistake costs scarce analyst time.

Mistakes in Other Domains

OCR	Spell Checker
Image Analysis	Human Eye
Translation	Low Expectation
Collaborative Filtering	Not much impact.

Mistakes in Other Domains

OCR	Spell Checker
Image Analysis	Human Eye
Translation	Low Expectation
Collaborative Filtering	Not much impact.

" [Recommendations are] guess work. Our error rate will always be high."
- Greg Linden (Amazon)

Building a Good Anomaly Detector

- Limit the detector's scope.
  - What *concrete* attack is the system to find?
  - Define a problem for which machine learning makes less mistakes.
- Gain insight into capabilities and limitations.
  - What exactly does it detect and *why*? What not and *why* not?
  - What are the features *conceptually* able to capture?
  - When exactly does it break?

  - Acknowledge shortcomings.
  - Examine false and true positives/negatives.

Image Analysis with Neural Networks

Tank

Image Analysis with Neural Networks

Tank

No Tank

What Can we Do?

- Limit the detector's scope.
  - What *concrete* attack is the system to find?
  - Define a problem for which machine learning makes less mistakes.
- Gain insight into capabilities and limitations.
  - What exactly does it detect and why? What not and why not?
  - What are the features conceptually able to capture?
  - When exactly does it break?
  - Acknowledge shortcomings.
  - Examine false and true positives/negatives.
- Assume the perspective of a network operator.
  - How does the detector help with operations?
  - Gold standard: work *with* operators. If they deem it useful, you got it right.

What Can we Do?

- Limit the detector's scope.
  - What *concrete* attack is the system to find?
  - Define a problem for which machine learning makes less mistakes.
- Gain insight into capabilities and limitations.

Once you have done all this ...
... you might notice that you now know enough about the activity you're looking for that you don't need any machine learning.

- Assume the perspective of a network operator.
  - How does the detector help with operations?
  - Gold standard: work *with* operators. If they deem it useful, you got it right.

Why is Anomaly Detection Hard?

The intrusion detection domain faces challenges that make it fundamentally different from other fields.

- Outlier detection and the high costs of errors
- Interpretation of results
- *Evaluation*
- *Training data*
- *Evasion risk*

Conclusion

- Machine learning for intrusion detection is challenging.
  - Reasonable and possible, but needs care.
  - Consider fundamental differences to other domains.
  - There is some good anomaly detection work out there.
- If you do anomaly detection, *understand* and *explain*.
- If you are given an anomaly detector, *ask questions*.

Conclusion

- Machine learning for intrusion detection is challenging.
  - Reasonable and possible, but needs care.
  - Consider fundamental differences to other domains.
  - There is some good anomaly detection work out there.
- If you do anomaly detection, *understand* and *explain*.
- If you are given an anomaly detector, *ask questions*.

“Open questions:

[...] Soundness of Approach: *Does the approach actually detect intrusions? Is it possible to distinguish anomalies related to intrusions from those related to other factors?”*

-Denning, 1987

Thanks for your attention.

Robin Sommer

International Computer Science Institute, &
Lawrence Berkeley National Laboratory

robin@icsi.berkeley.edu
http://www.icir.org

Thanks for your attention.

Robin Sommer

International Computer Science Institute, &

Lawrence Berkeley National Laboratory

robin@icsi.berkeley.edu

http://www.icir.org