Support Vector Machines for Intrusion Detection

Invited Talk at Tsinghua University,
organized by Research Institute of Information Technology

Dr. Tran Quang Anh

CERNET Computer Emergency and Response Team

Email: qa_at_cernet.edu.cn

22, Apr. 2004

Contents

- Intrusion Detection
- Support Vector Machines
- SVM for Flow-based Intrusion Detection
- One-class SVM for Time-based Intrusion Detection
- Conclusions

Intrusion Detection

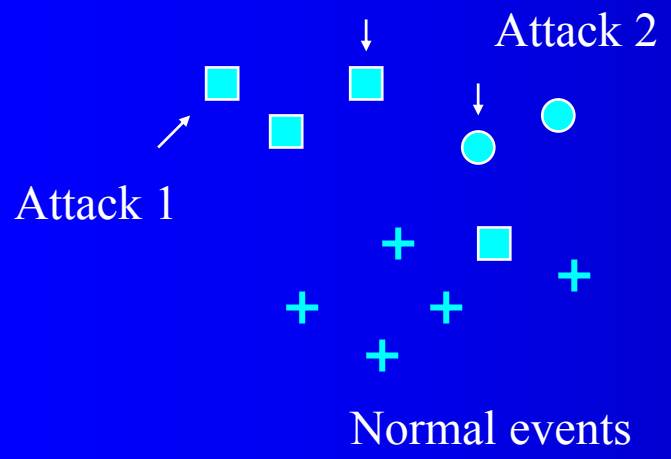
- The PDRR model for computer network security
  - Protection
  - Detection
  - Response
  - Recovery

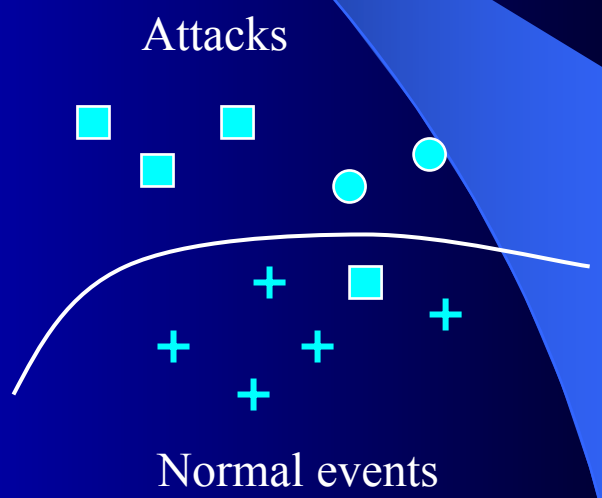
Intrusion Detection

- Common Intrusion Detection Framework
  - Event generators
    - Host-based
    - Network-based
  - Event analyzers
    - Misuse detection
    - Anomaly detection
  - Event databases
  - Response units

Intrusion Detection

- On the traditional view



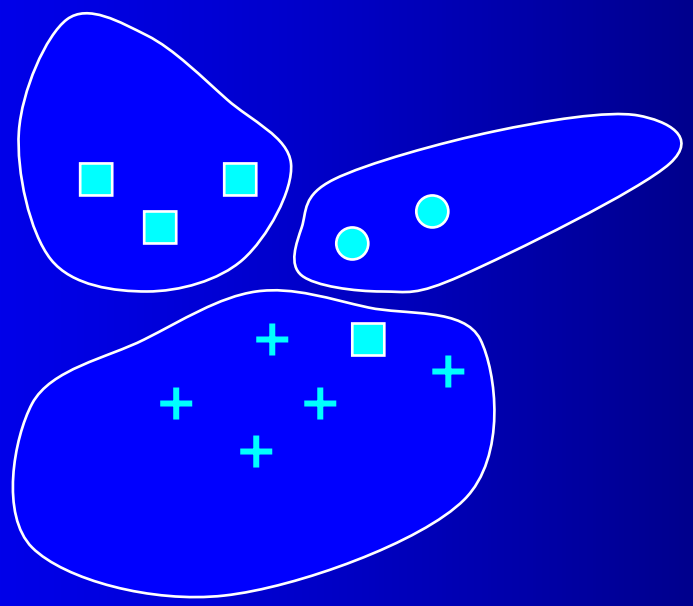


Misuse detection

Anomaly detection

Intrusion Detection

- One the classification-based view



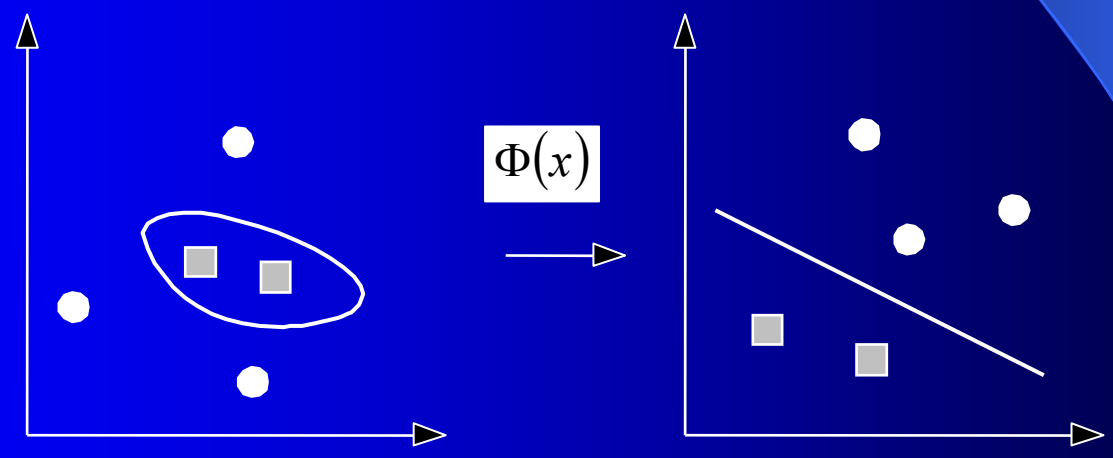
Misuse detection and anomaly detection

Support Vector Machines

- Key techniques
  - Kernel method
  - Maximal margin hyperplane
  - Trading off empirical risk and complexity

Support Vector Machines

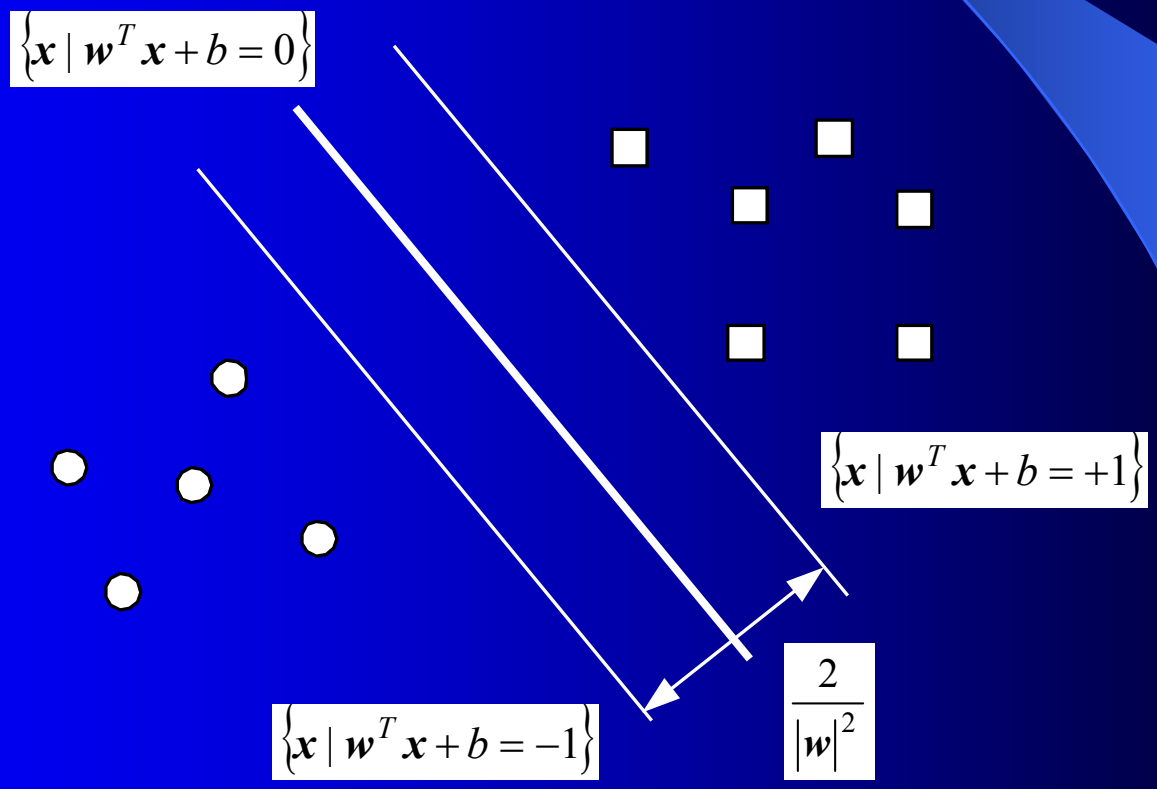
- Kernel method



$$K(\mathbf{x}_i, \mathbf{x}_j) = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_j)$$

Support Vector Machines

- Maximal margin hyperplane



Support Vector Machines

- Trading off empirical risk and complexity
  - C-SVC (Cortes, Vapnik 95)

    $C, C_+, C_- > 0$
  - ν-SVC (Schoelkopf 98)

    $0 \le \nu \le 1$ $\quad C = 1/\rho$
  - One-class SVM (Schoelkopf 99)

    $0 \le \nu \le 1$

SVM for Flow-based IDS

- Event generator
  - DARPA intrusion detection evaluation 98-99
  - SIG-KDD 99
  - Flow based
  - 41 features: TCP/IP (1-9); Host log (10-22); Time-based traffic statistics (23-31); Event-based traffic statistics (32-41)
  - R2L, U2R, DOS, PROBE, NORMAL

SVM for Flow-based IDS

- Event analyzer
  - C-SVC
  - SVM*light*

SVM for Flow-based IDS

- Feature and training model selection
  - S: Feature model
  - M: Training model
  - G: Generalization performance
  - $\max_{S,M} G(S,M)$
  - $\xi\alpha$ - estimate (Joachims 99)
  - Genetic Algorithms

SVM for Flow-based IDS

- Improve the detection speed
  - Reduce the number of support vectors
  - Clustering technique

SVM for Flow-based IDS

- Detection performance

	NORMAL	PROBE	DOS	U2L	R2L
NORMAL	0	1	2	2	2
PROBE	1	0	2	2	2
DOS	2	1	0	2	2
U2L	3	2	2	0	2
R2L	4	2	2	2	0

Bagged Boosting	0.2331
Kernel Miner	0.2474
MP13	0.2552
E-SVM	**0.2412**

One-class SVM for Time-based IDS

- Event generator
  - DARPA intrusion detection evaluation 98-99
  - Network traffic statistics on certain duration
  - TCPSTAT
  - # of ICMP, TCP, UDP, average packer size, deviation of packet size
  - Portsweep, mailbomb, ipsweep, satan, neptune
  - Training set – first week data (attack free)
  - Testing set – second week data

One-class SVM for Time-based IDS

- Event analyzer
  - One-class SVM
  - LIBSVM

One-class SVM for Time-based IDS

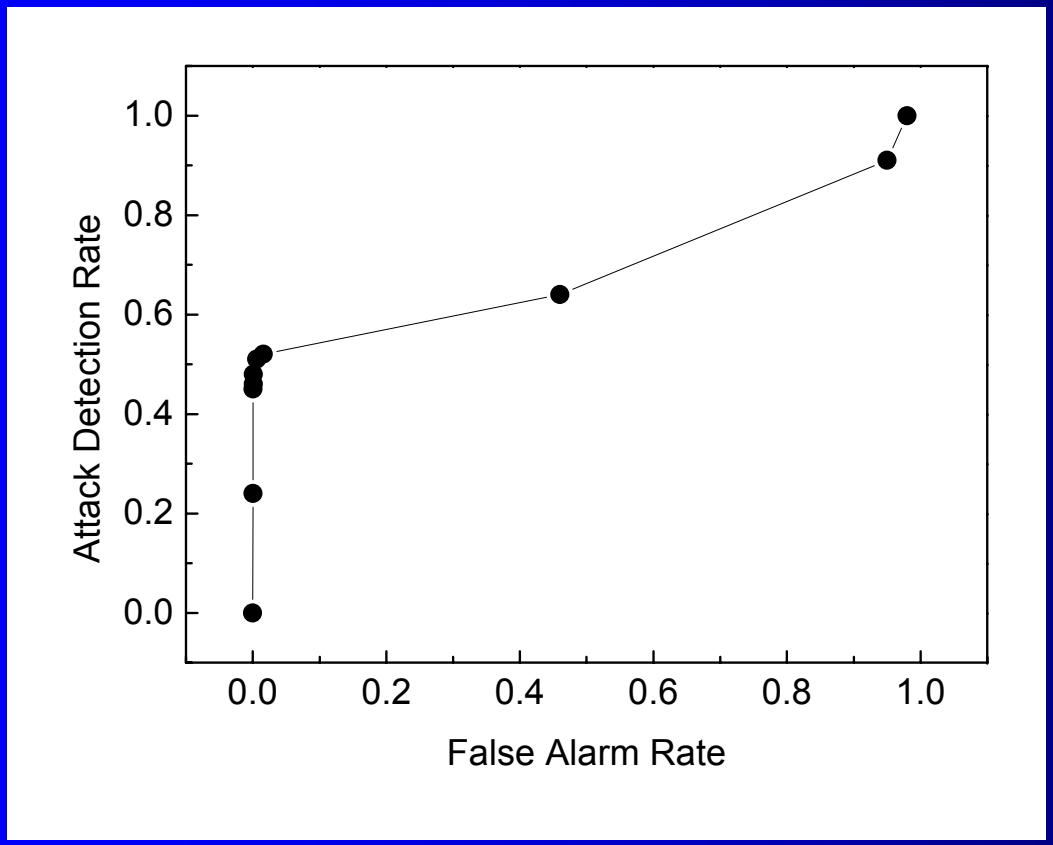
- Feature and training model selection
  - M: Training model
  - G: Generalization performance
  - $\max_{M} G(M)$
  - $\xi\alpha\rho$ - estimate (Tran 03)
  - Genetic Algorithms

One-class SVM for Time-based IDS

- Improve the detection speed
  - Reduce the number of support vectors
  - Clustering technique

One-class SVM for Time-based IDS

- Detection performance – ROC curve



Duration = 60s

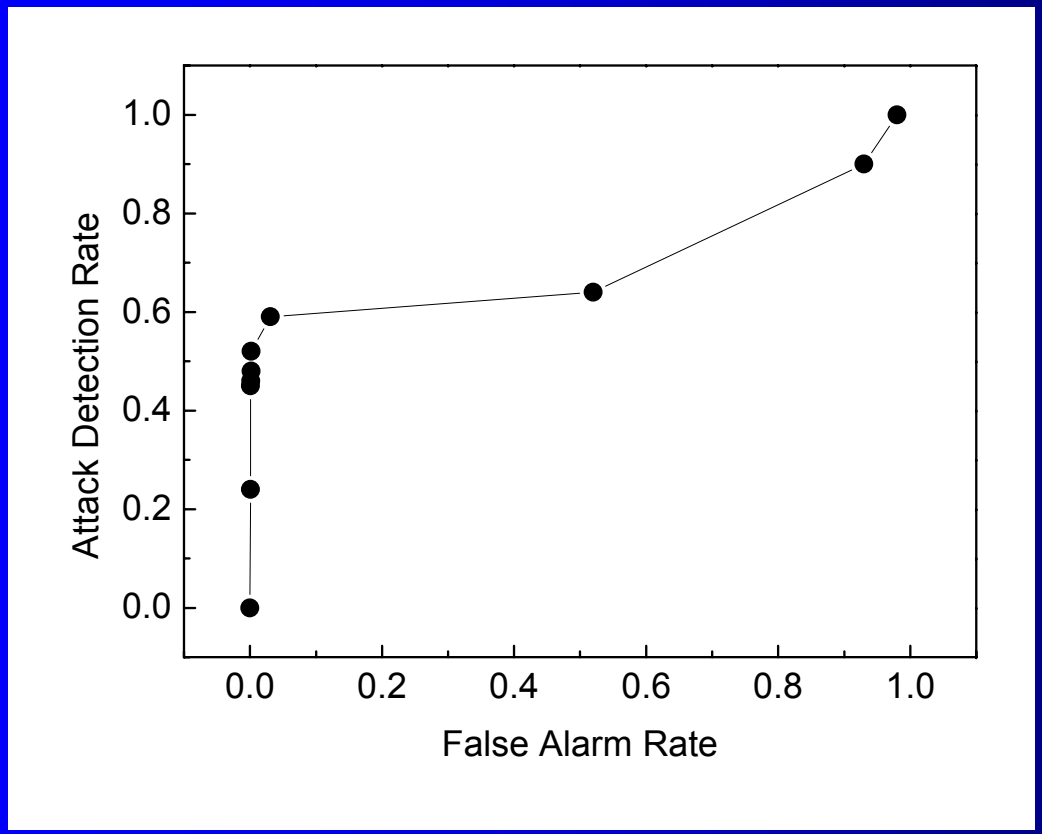
Training size = 6601

Testing size = 6424

attack = 137

One-class SVM for Time-based IDS

- Detection performance – ROC curve



Duration = 120s

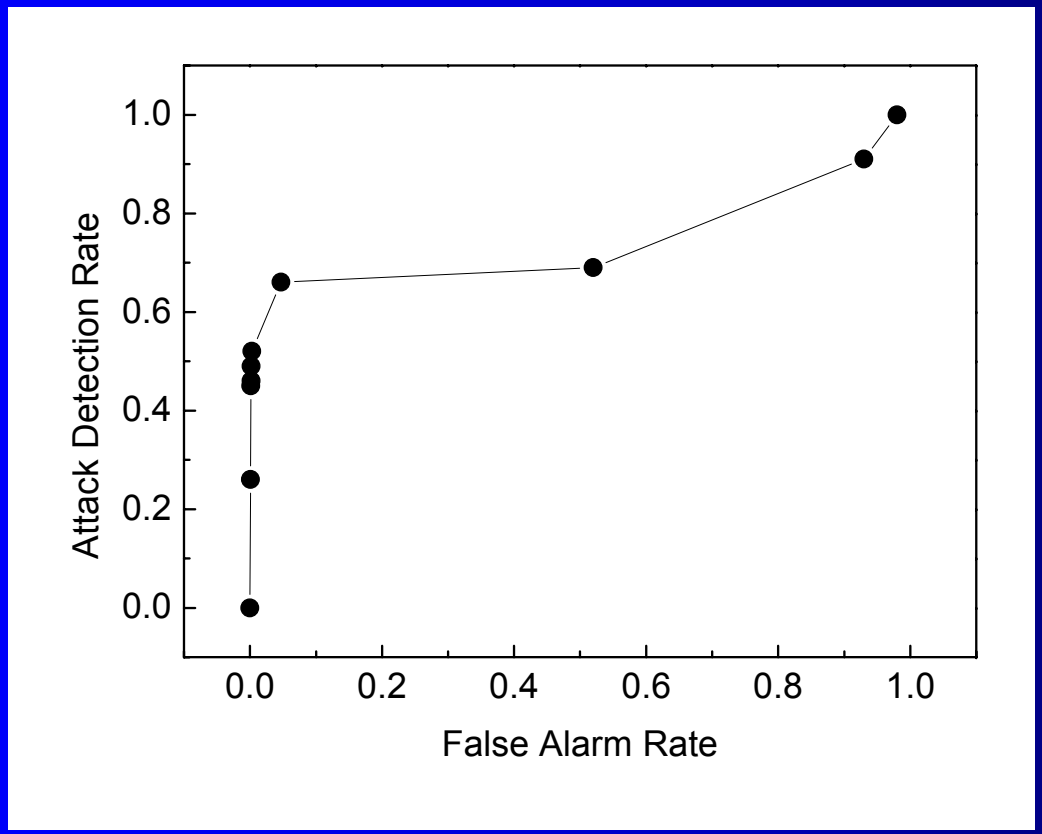
Training size = 3301

Testing size = 3212

attack = 75

One-class SVM for Time-based IDS

- Detection performance – ROC curve



Duration = 180s

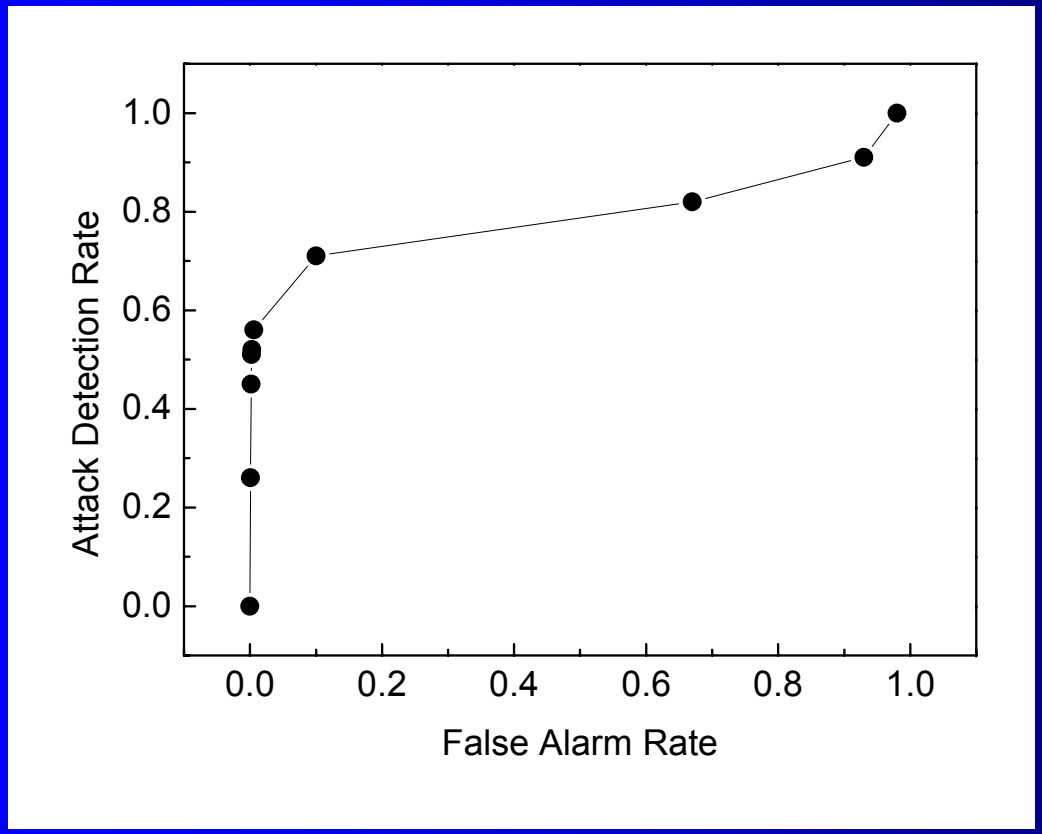
Training size = 2201

Testing size = 2142

attack = 53

One-class SVM for Time-based IDS

- Detection performance – ROC curve



Duration = 300s

Training size = 1321

Testing size = 1285

attack = 37

One-class SVM for Time-based IDS

- Detection performance at 100 false alarms

	60s	120s	180s	300s
Detection rate	0.52	0.59	0.66	0.71
False alarm	0.016	0.031	0.047	0.1

Conclusions

- Attack feature: specific patterns and statistics
- Network data: flow-based and time-based
- Analyzer: C-SVM and one-class SVM
- Improve detection speed: clustering techniques
- Feature and training model selection: evolving algorithms

References

- Porras P., The Common Intrusion Detection Framework Architecture, URL: http://gost.isi.edu/cidf/drafts/architecture.txt
- Vapnik V.N.. 1999. An overview of statistical learning theory, IEEE Trans. on Neural Networks, vol. 10, Issue 5, pp. 988 -999, Sept. 1999
- Schölkopf, B., Platt, J, et al. 2001. Estimating the support of a high-dimensional distribution. Neural Computation, 13, 2001, 1443-1471
- Joachims T. Estimating the generalization performance of a SVM efficiently. Proceedings of the Seventeenth International Conference on Machine Learning, San Francisco, 2000
- DARPA Intrusion Detection Evaluation, URL: http://www.ll.mit.edu/IST/ideval/
- KDD Cup 1999 Data. URL: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html
- LIBSVM, URL: http://www.csie.ntu.edu.tw/~cjlin/libsvm/
- SVMlight, URL: http://svmlight.joachims.org
- SNORT, URL: http://www.snort.org

Thank you!

Tran Quang Anh

22 Apr 2004