Apprentissage profond pour la détection d'anomalies dans un environnement hôte et la cybersécurité

Wahab Hamou-Lhadj, PhD
Université Concordia
wahab.hamou-lhadj@Concordia.ca

87e Congrès ACFAS
UQO, Gatineau, QC, Canada
30 mai 2019

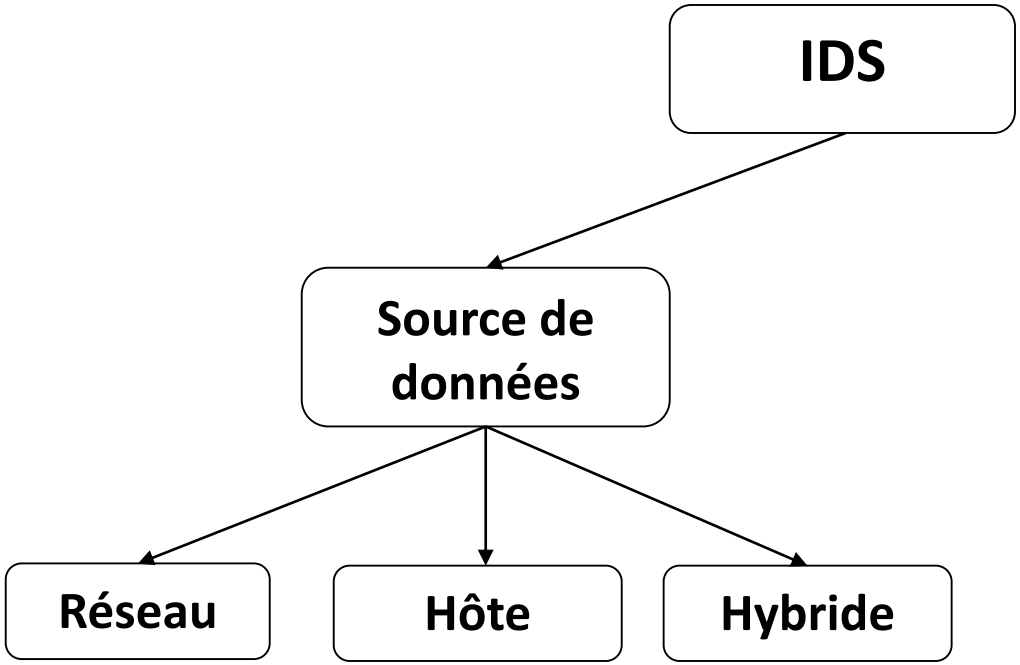
Système de détection d'intrusions (IDS)

- Un système de détection d'intrusions (IDS) est un dispositif logiciel ou matériel ou une combinaison des deux.

- Chargé de surveiller un réseau ou un hôte donné afin de prévoir ou d' identifier toute action suspecte et non-autorisée et éventuellement réagir à cette action.

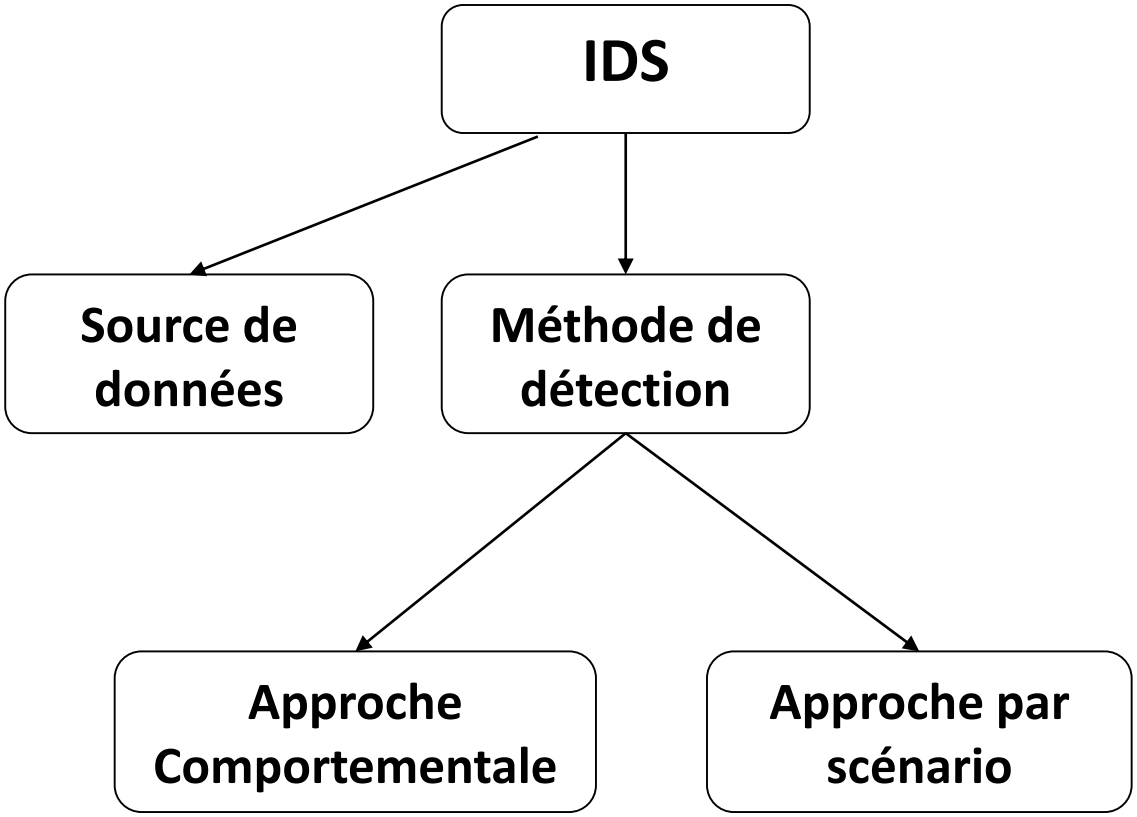
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Classification des IDS



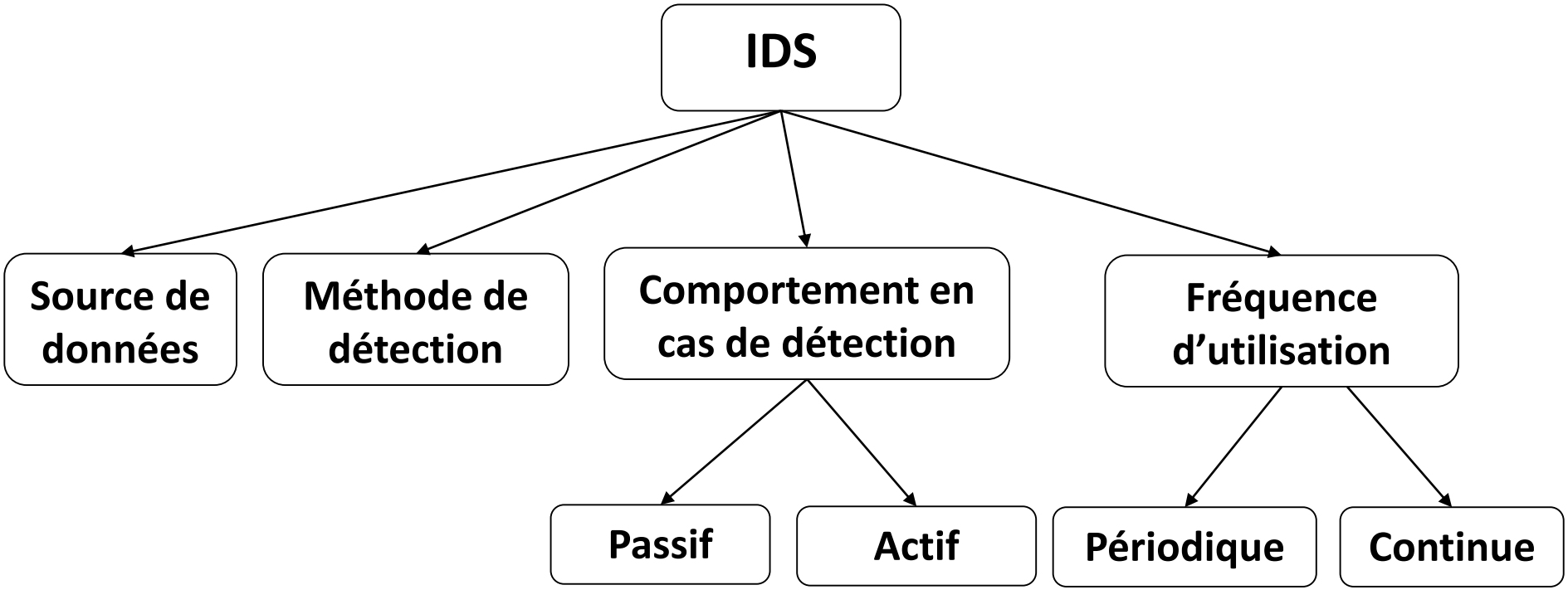
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Classification des IDS



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Classification des IDS



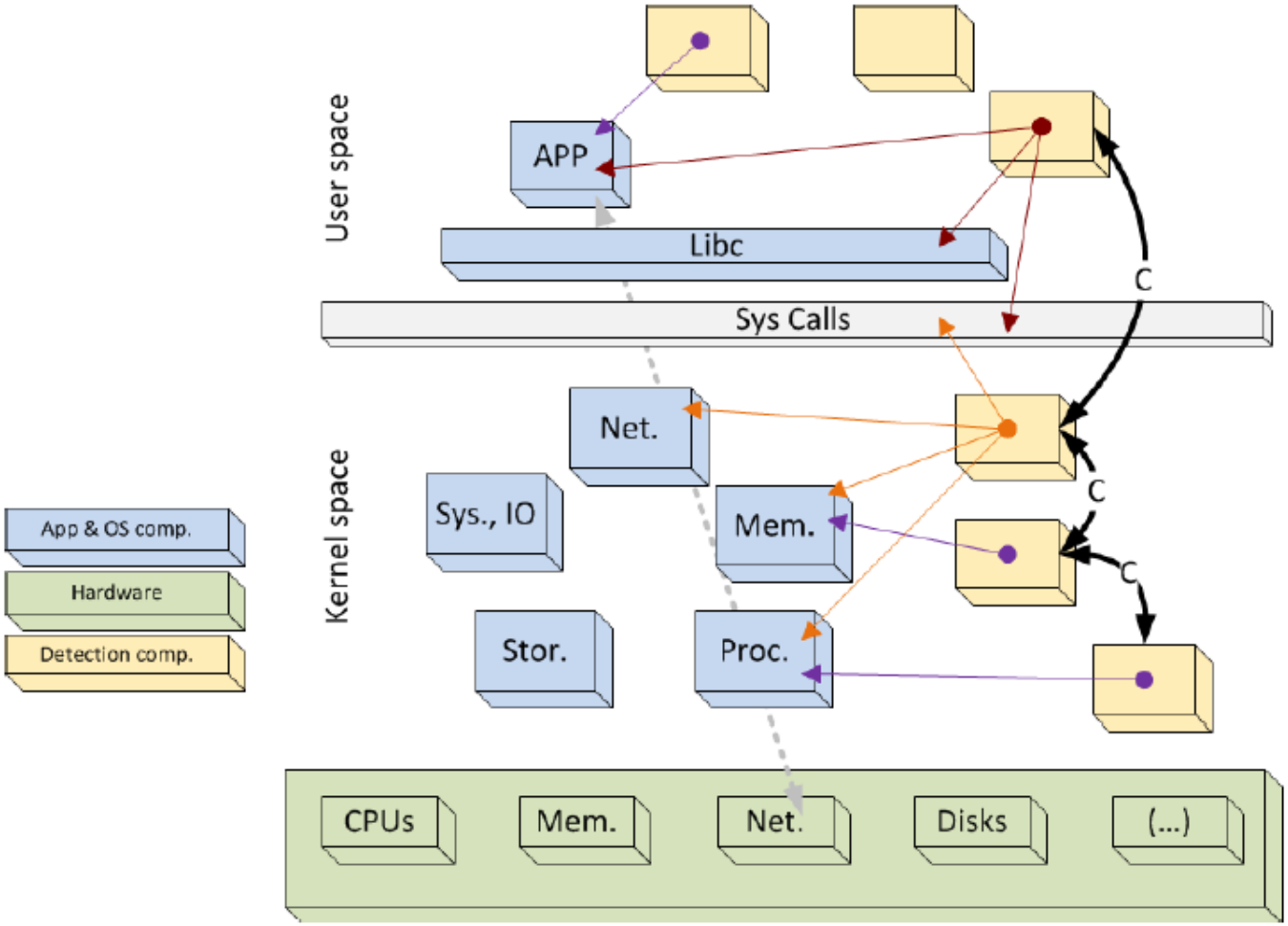
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Système de détection d'anomalies au niveau de hôte (HADS)

- Par scénario:
  - + Prise en compte des caractéristiques exactes des attaques
  - – Base de signatures difficiles à construire et mettre à jour
  - – Pas de détection d'attaques inconnues
- Analyse comportementale
  - + Permet de détecter des nouvelles attaques
  - – Difficile de modéliser le comportement normal
  - – Taux élevé de faux positifs

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

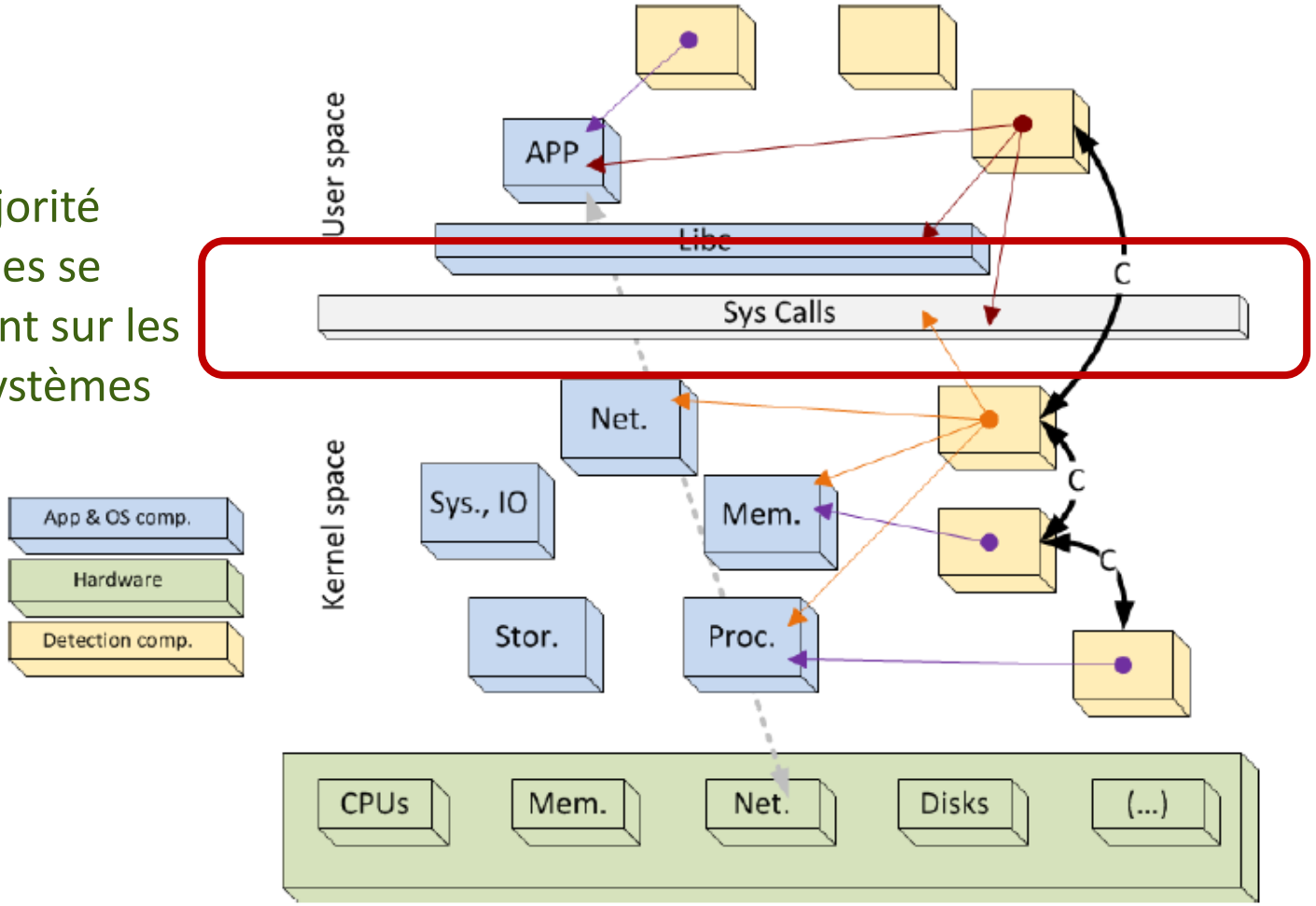
Monitoring dans un environnement hôte - vue simplifiée



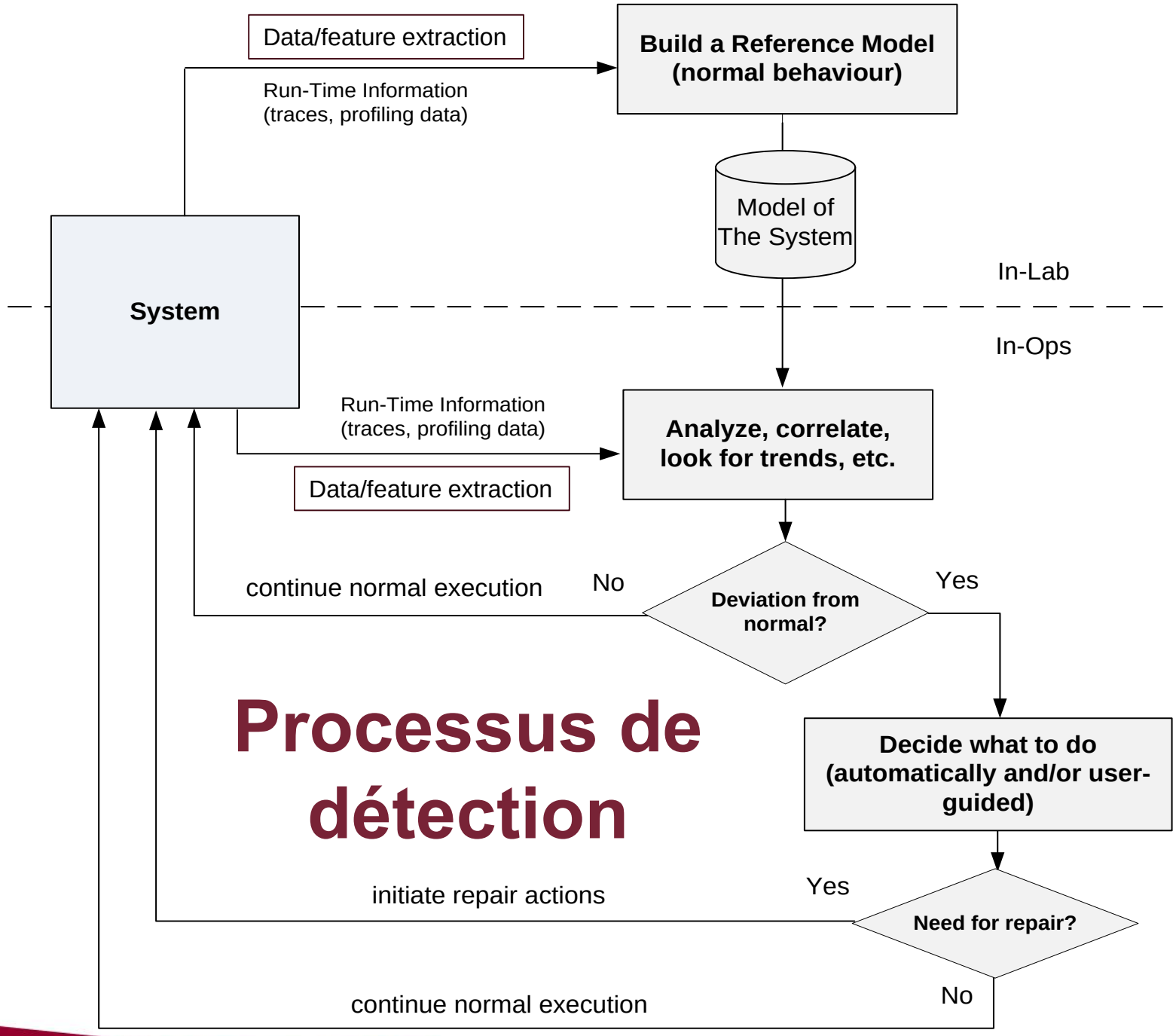
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Monitoring dans un environnement hôte - vue simplifiée

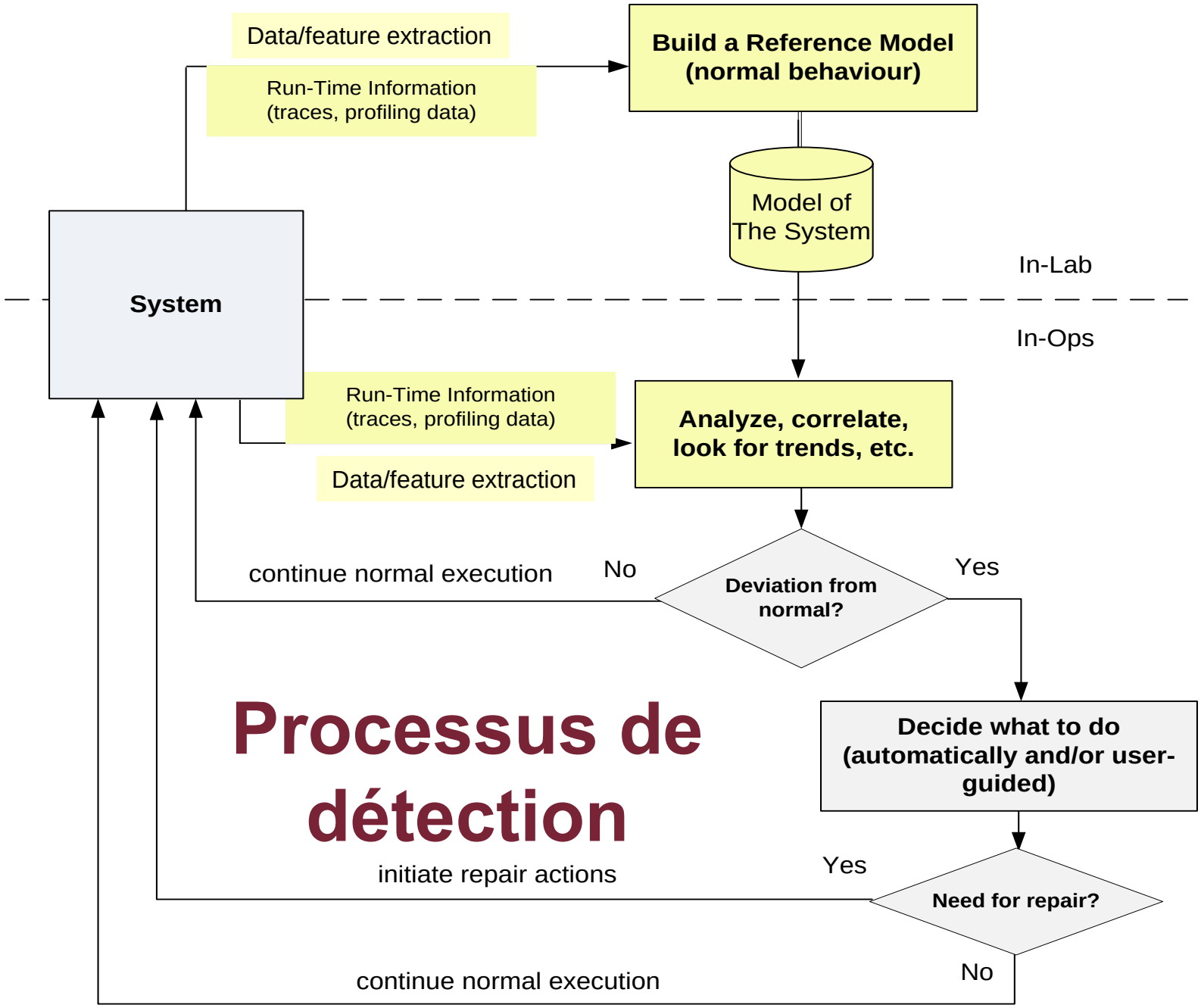
La majorité d'études se concentrent sur les appels systèmes



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)



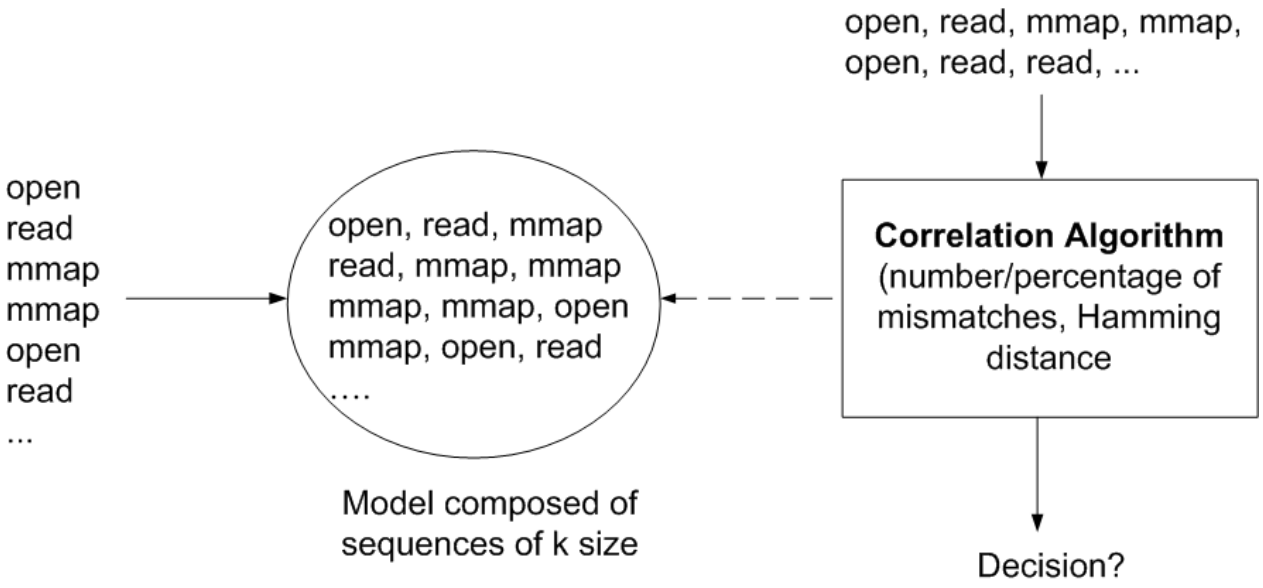
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Techniques basées sur l'apprentissage machine

- Plusieurs techniques ont été utilisées pour modéliser le comportement normal d'un système
  - STIDE, HMM, ANN, KNN, SVM, RF, modèles statistiques, etc.



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Techniques basées sur l'apprentissage profond

- Une famille d'algorithmes ML devenue populaire grâce aux progrès des capacités de calcul et les infrastructures BigData

- Applications:
  - La classification d'images
  - Les modèles génératifs
  - La reconnaissance de la parole
  - Le traitement automatique du langage naturel
  - La traduction automatique

- Les études préliminaires montrent que l'apprentissage profond fournit de meilleurs résultats pour le HADS et le NIDS

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Plusieurs types d'algorithmes

- Les réseaux de croyance profonde
- Les auto-encodeurs
- Les réseaux de neurones à convolution
- Les réseaux de neurones récurrents
- Combinaison de ces techniques

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Les réseaux de neurones récurrents (RNN)

- Les RNN opèrent sur des données séquentielles
- Adéquats pour la détection d'anomalies basée sur les appels systèmes
  - Les informations séquentielles sont conservées dans les états cachés du réseau de neurones récurrent (on parle de mémoire)
  - Les RNN permettent de trouver une corrélation entre les évènements (dépendances à long terme)
- Deux variantes: LSTM (Long Short Term Memory) et GRU (Gated Reccurent Units)

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Revue de la littérature

DataSet	IDS	Description	Type	References
CTU-UNB	NIDS	CTU-UNB ucs [2017] dataset consists of various botnet traffics from CTU-13 dataset [20] and normal traffics from the UNB ISCX IDS 2012 dataset Shiravi et al. [2012]	Hexadecimal	Yu et al. [2017]
Contagio-CTU-UNB	NIDS	Contagio-CTU-UNB dataset con-	Text	Yu et al. [2017].

DEEP LEARNING FOR ANOMALY DETECTION: A SURVEY

Raghavendra Chalapathy
University of Sydney,
Capital Markets Co-operative Research Centre (CMCRC)
rcha9612@uni.sydney.edu.au

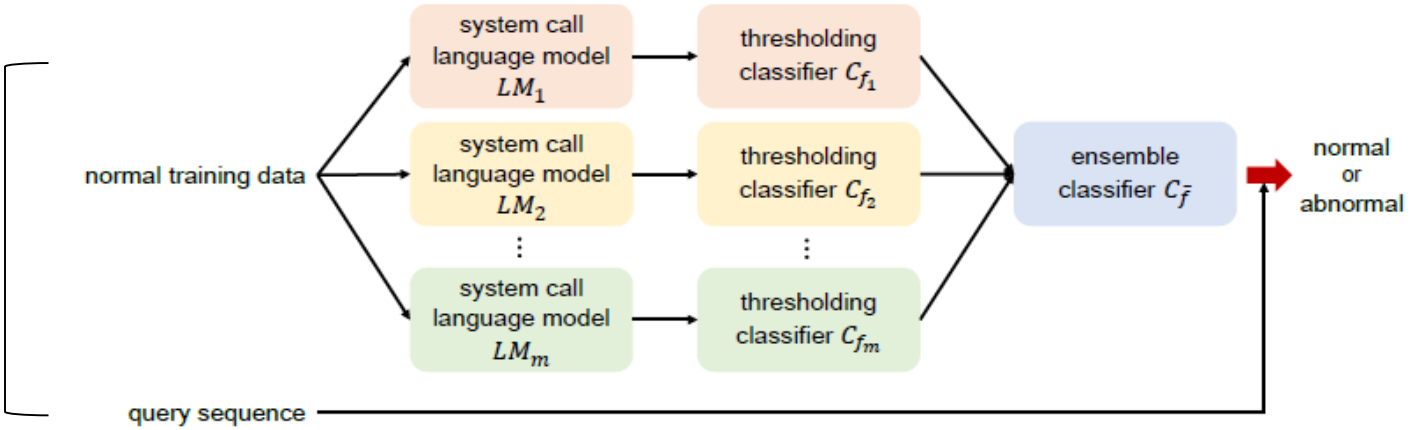
Sanjay Chawla
Qatar Computing Research Institute (QCRI), HBKU
schawla@qf.org.qa

January 24, 2019

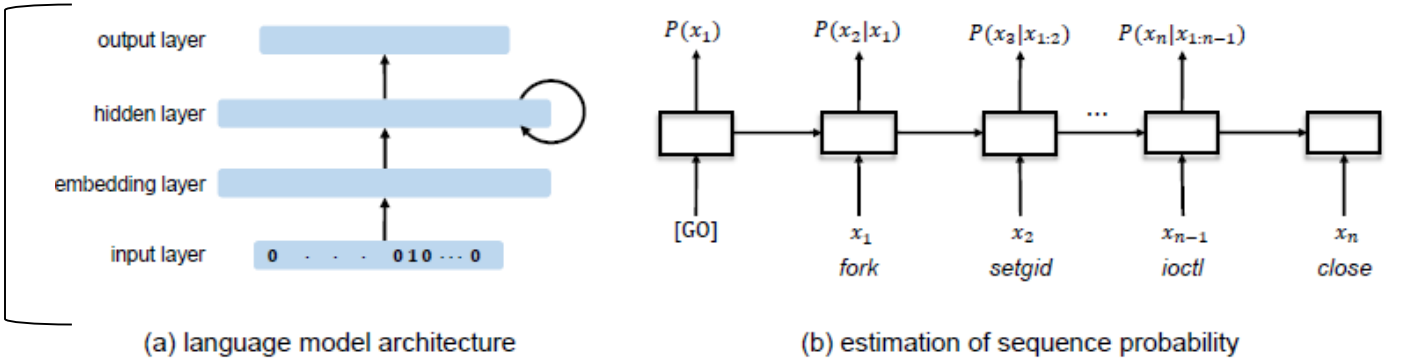
		fic capturedfrom backbone links between Japan and USA. Every daysince 2007		
Realistic Global Cyber Environment (RGCE)	NIDS	RGCE jam [2009] contains realistic Internet Service Providers (ISPs) and numerous different web services as in the real Internet.	Text	Zolotukhin et al. [2016]
ADFA-LD	HIDS	The ADFA Linux Dataset (ADFA-LD). This dataset provides a contemporary Linux dataset for evaluation by traditional HIDS Creech and Hu [2014]	Text	Kim et al. [2016], Chawla et al. [2018]
UNM-LPR	HIDS	Consists of system calls to evalute HIDS system University [2012]	Text	Kim et al. [2016]
Infected PDF samples	HIDS	Consists of set of Infected PDF samples, which are used to monitor the malicious traffic	Text	Chen et al. [2018]

Modélisation du langage des appels système basée sur LSTM

2. Minimisation des faux positifs en combinant plusieurs classifieurs



1. Modélisation du langage appels système

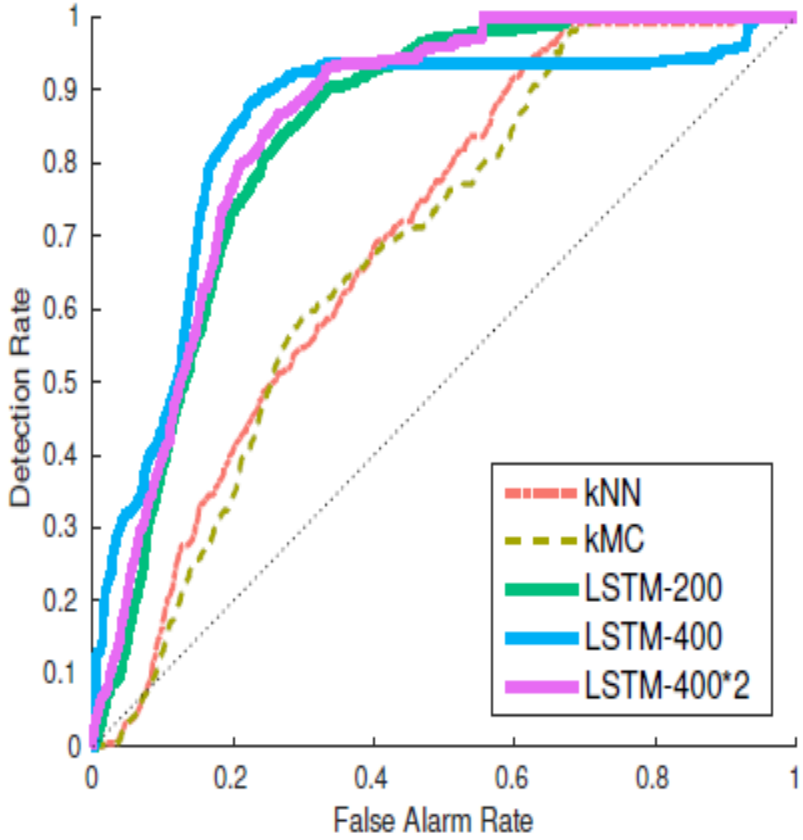


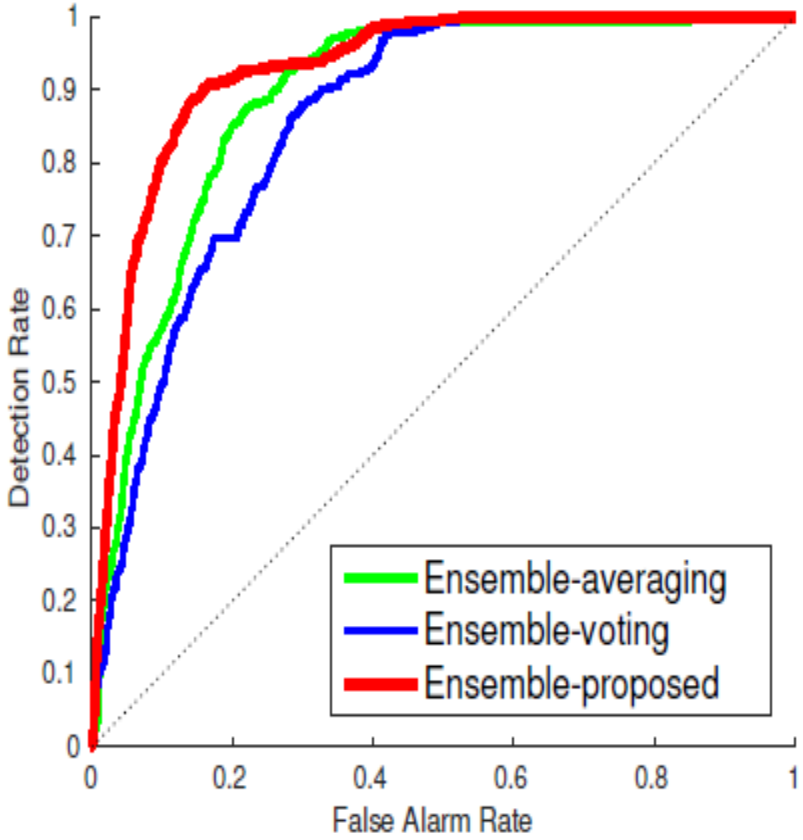
(a) language model architecture (b) estimation of sequence probability

Architecture proposée par Kim, G., Yi, H., Lee, J., Paek, Y., & Yoon, S. (2016). https://arxiv.org/abs/1611.01726

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Modélisation du langage des appels système basée sur LSTM





Architecture proposée par Kim, G., Yi, H., Lee, J., Paek, Y., & Yoon, S. (2016). https://arxiv.org/abs/1611.01726

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Comparaison

Architecture	Modèle	Données	AUC
LSTM (Kim et al., 2016)	LSTM et modèle de langage des appels système	ADFA-LD KDD UNM	0.92 0.99 0.96
CNN/RNN (Chawla et al., 2018)	Une couche à 200 unités de GRU Une couche à 200 unités de LSTM 6 couches de 1D CNN, 200 unités de GRU 7 couches de 1D CNN, 500 unités de GRU 8 couches de 1D CNN, 600 unités de GRU	ADFA-LD	0.66 0.74 0.80 0.79 0.81
GRU (Lv et al., 2018)	Modèle de prédiction des séquences (approche séquence à séquence)	ADFA-LD	0.96

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Projet de surveillance avancée dans un environnement hôte

- Projet NSERC-DND de 4 ans (2012-2016)
- 6 PhDs, 8 Maîtrises, 2 Postdocs, 2 ARs

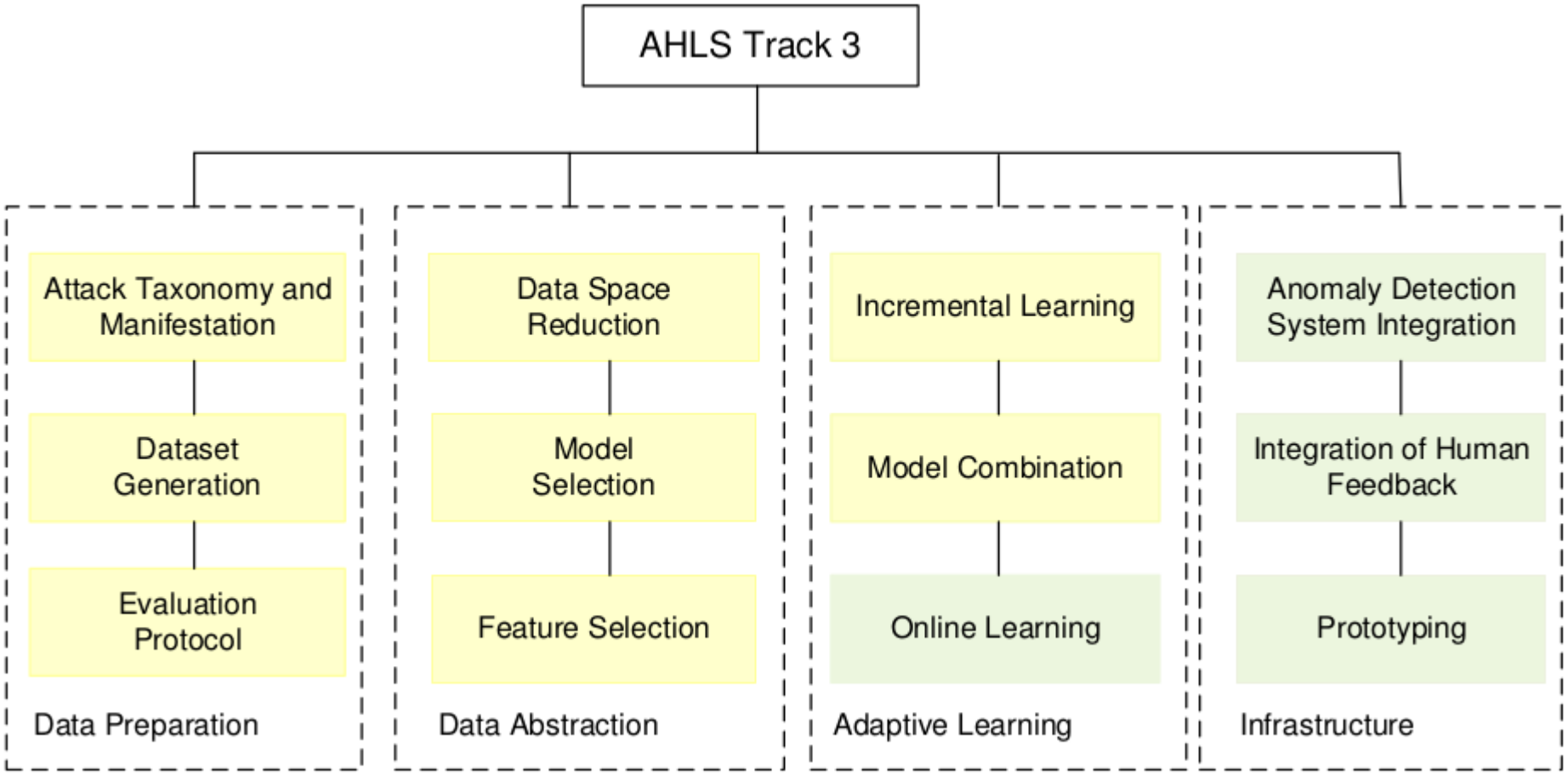
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Objectifs du projet

- Monitorer et protéger les systèmes hôtes
- Développer des HADS modulaires, adaptables et évolutifs
- Réduire les faux positifs et améliorer les vrais positifs
- Développer des testbeds complets et des protocoles d'évaluation
- Fournir des recommandations pour des recherches et orientations futures

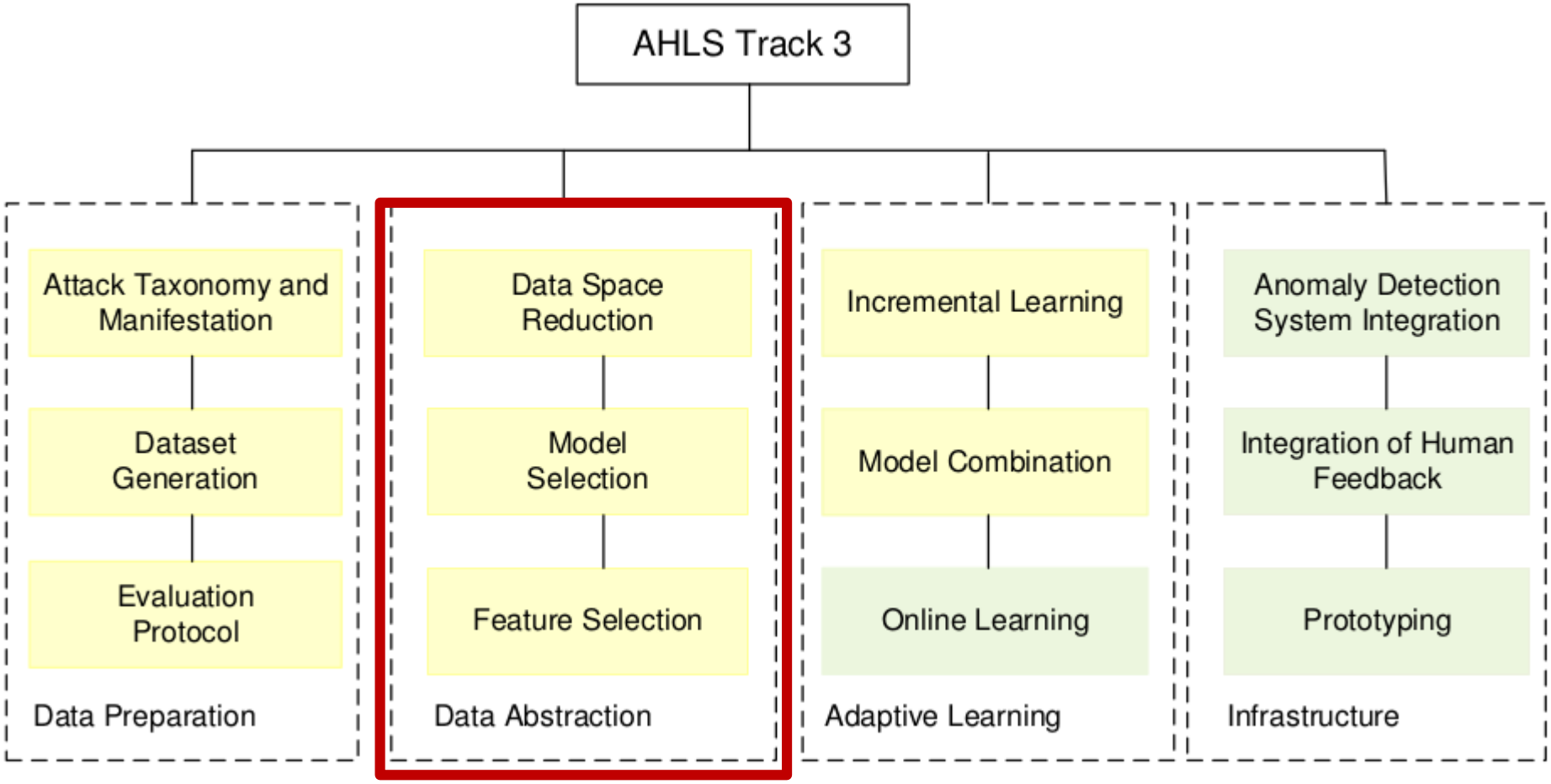
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Structure du projet



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Structure du projet



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

La méthode KSM (Kernel State Modeling)

- KSM est une technique de détection d'anomalies
  - Transforme les appels système en modules du noyau, appelés états
  - Détecte les anomalies au niveau de l'interaction des états du noyau
  - Réduit l'espace de données utilisé pour la formation et les tests
  - Favorise l'efficacité tout en conservant la précision

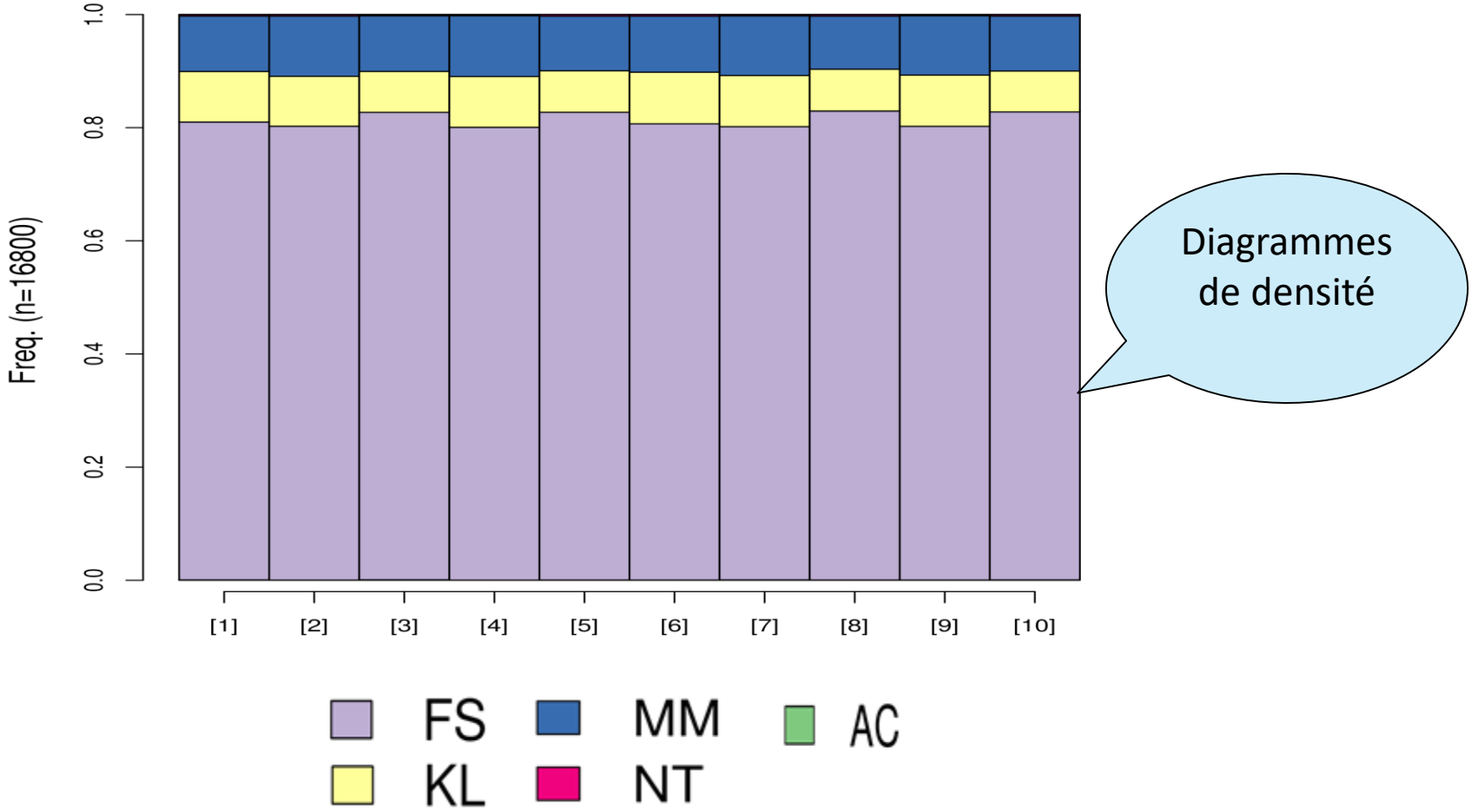
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Transformation d'appels système en états de modules de noyau

State	Module in Linux Source Code	# of System Calls
AC	Architecture	10
FS	File System	131
IPC	Inter Process Communication	7
KL	Kernel	127
MM	Memory Management	21
NT	Networking	2
SC	Security	3
UN	Unknown	37

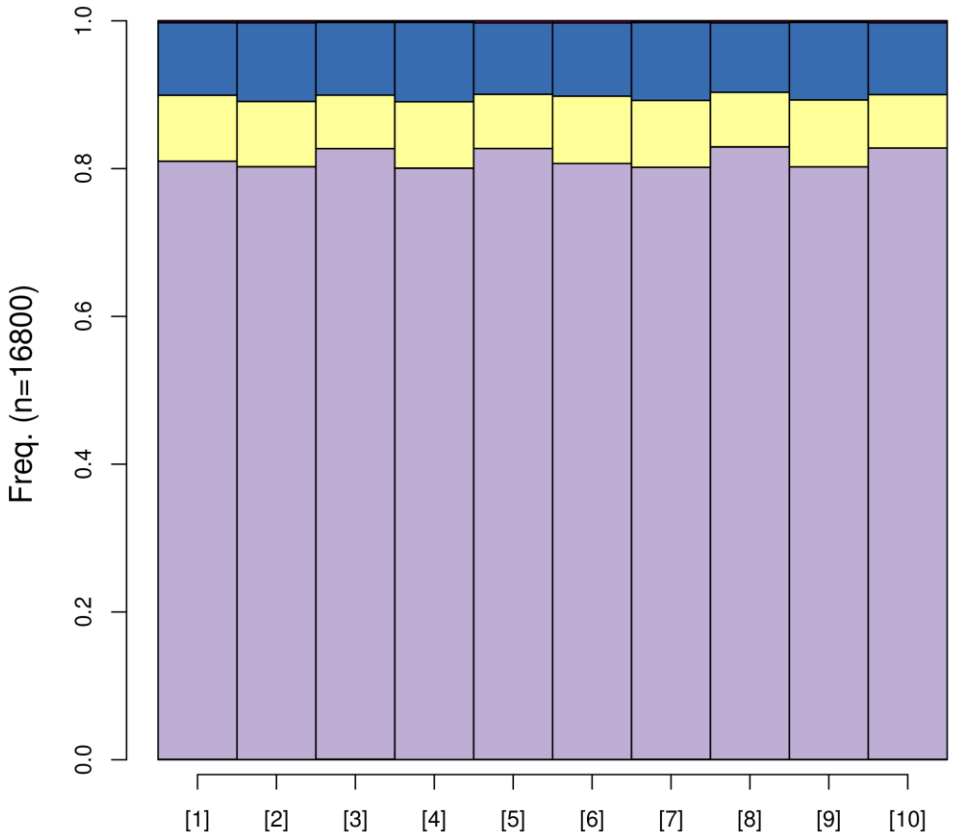
Source: http://syscalls.kernelgork.com

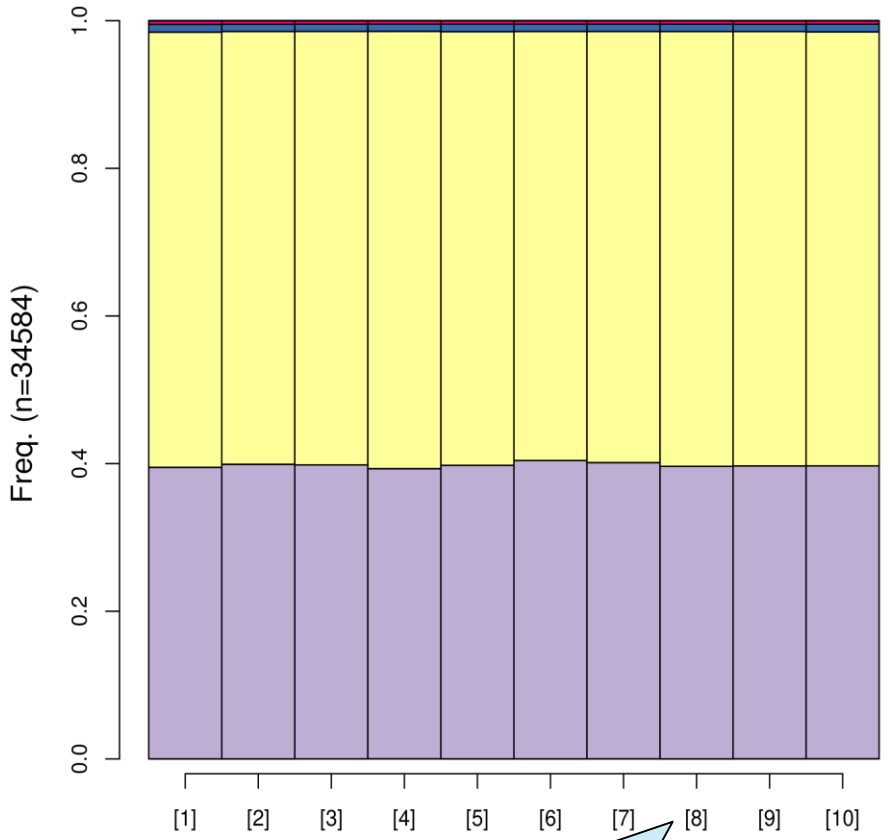
KSM et diagrammes de densité



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

KSM et diagrammes de densité





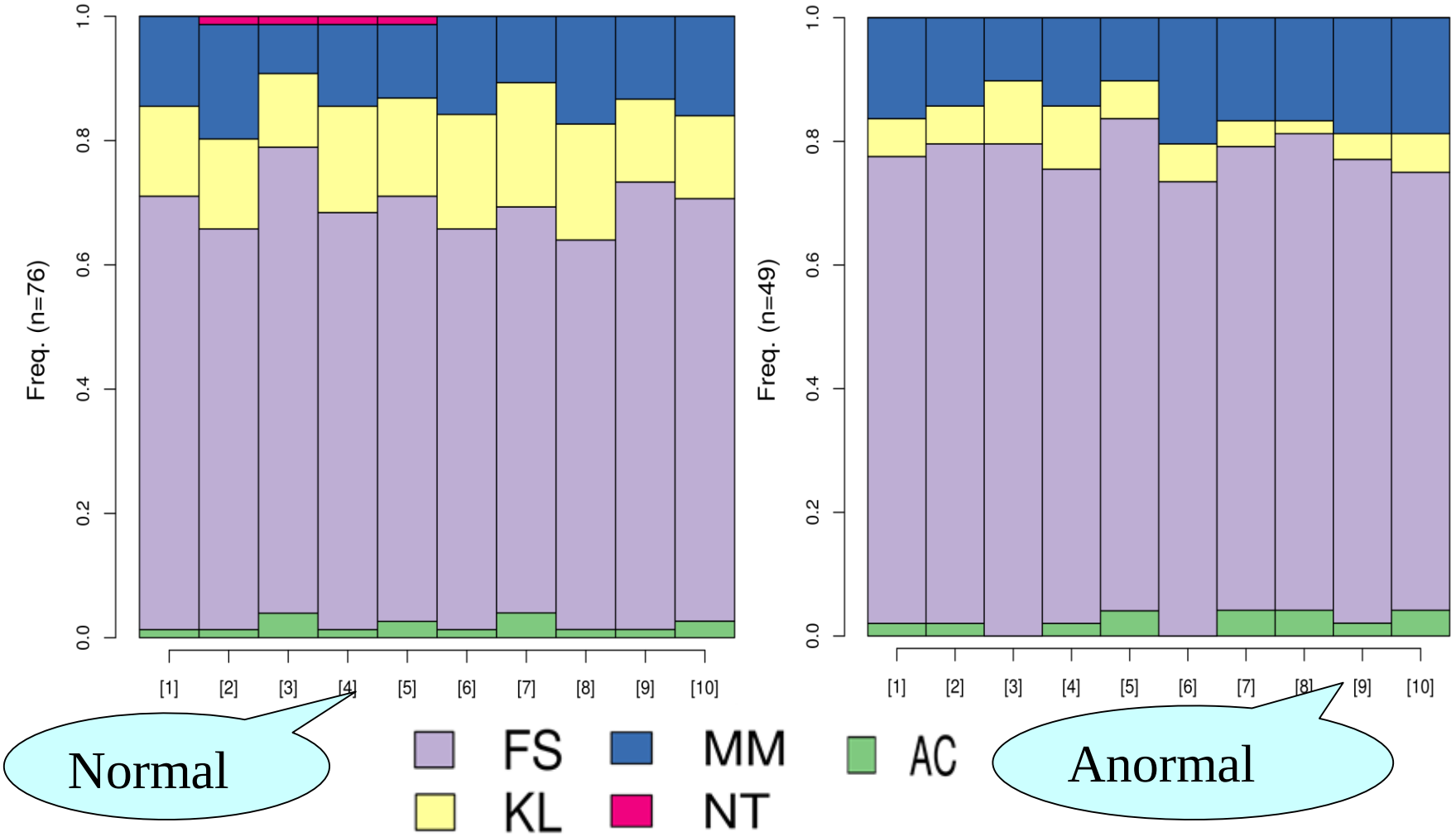
Normal

FS MM AC
KL NT

Anormal

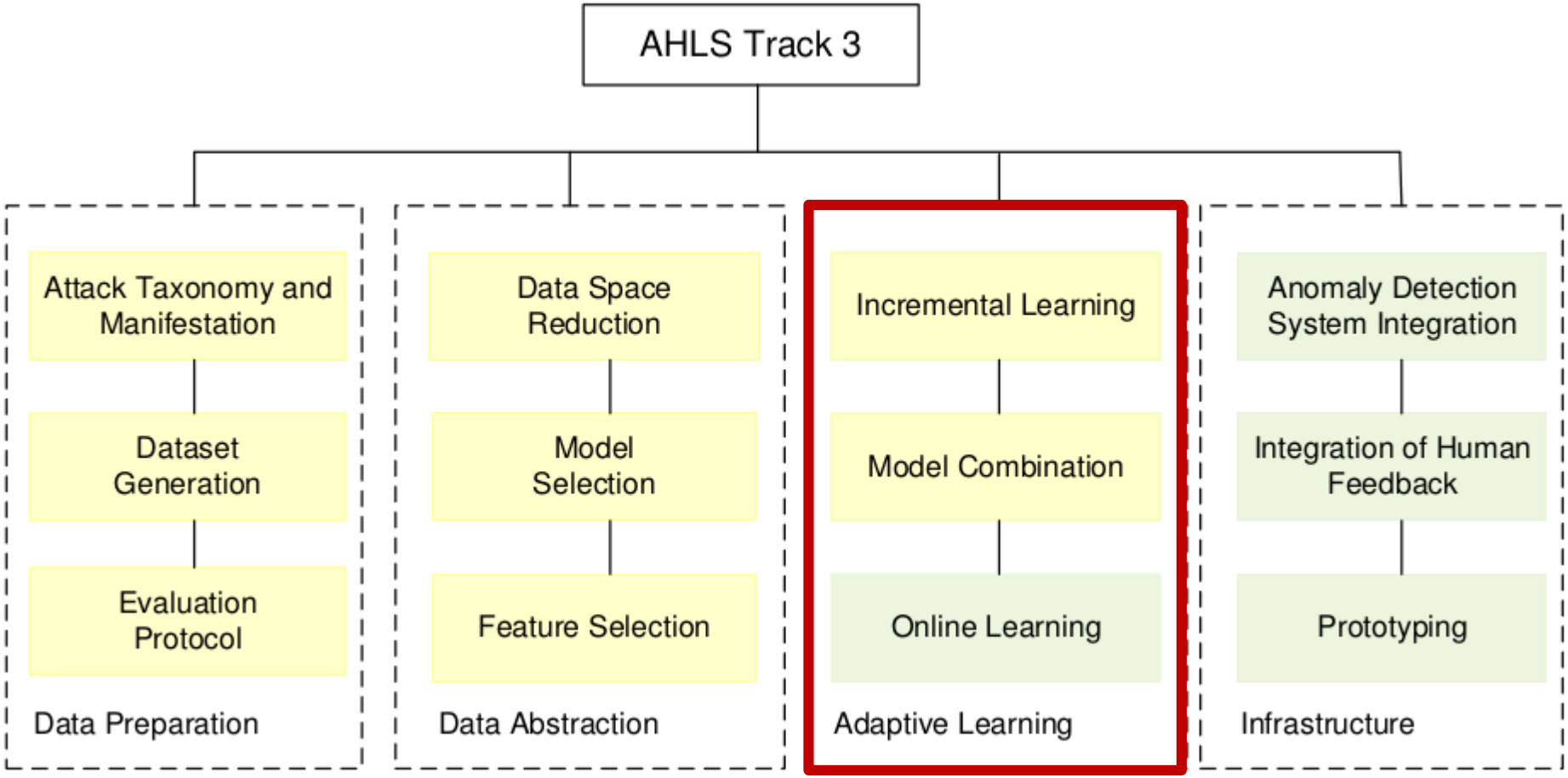
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

KSM et diagrammes de densité



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Structure du projet



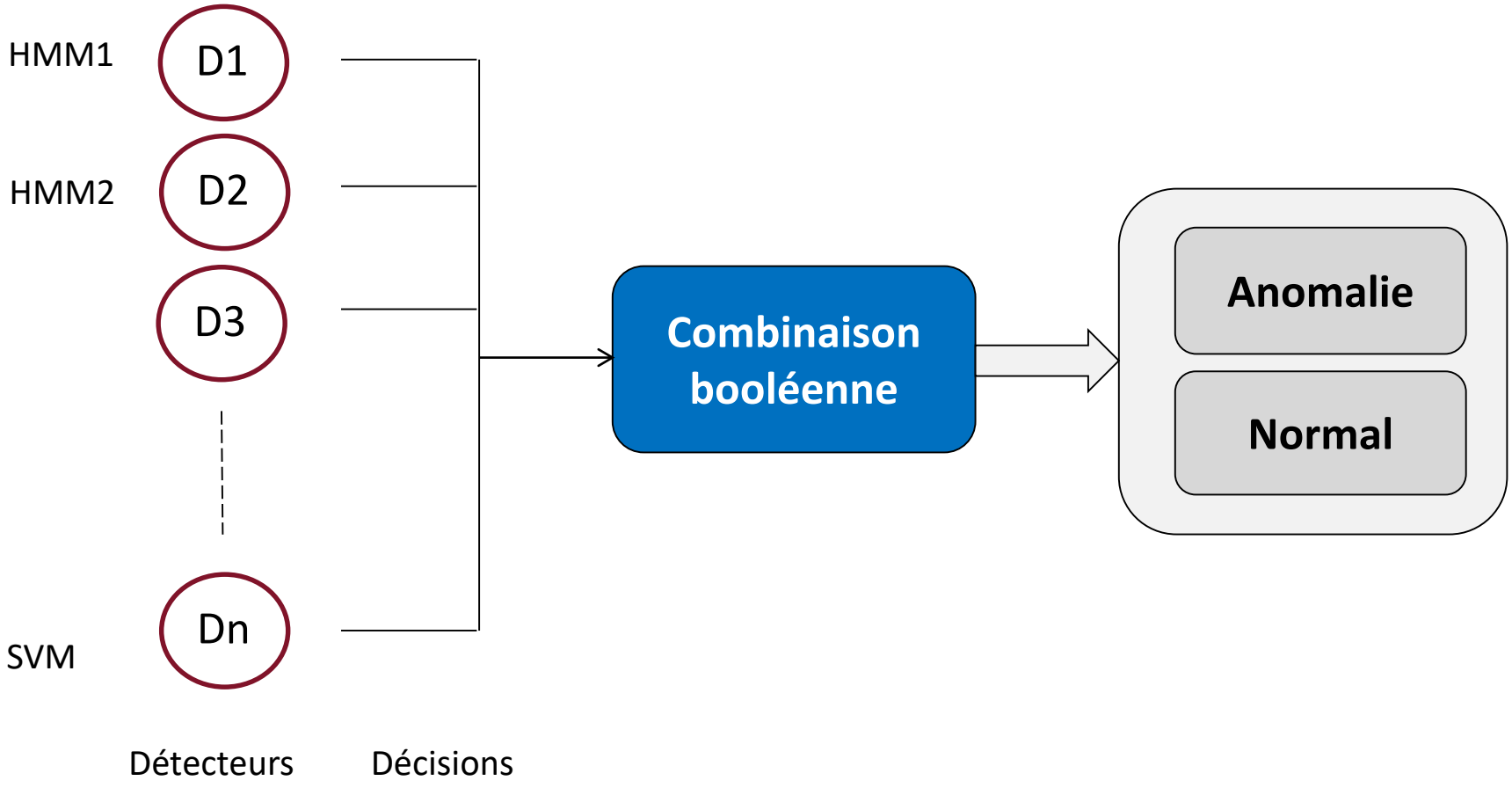
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Combinaison des modèles

- Un seul classifieur peut ne pas fournir une bonne approximation de la structure ou de la distribution des données
  - Pas de classifieur dominant pour toutes les distributions des données (théorème «pas de repas gratuit»)
  - La vraie distribution des données est généralement inconnue
  - Une quantité limitée de données (étiquetées) est généralement fournie pendant l'apprentissage

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Combinaison de réponses de différents détecteurs



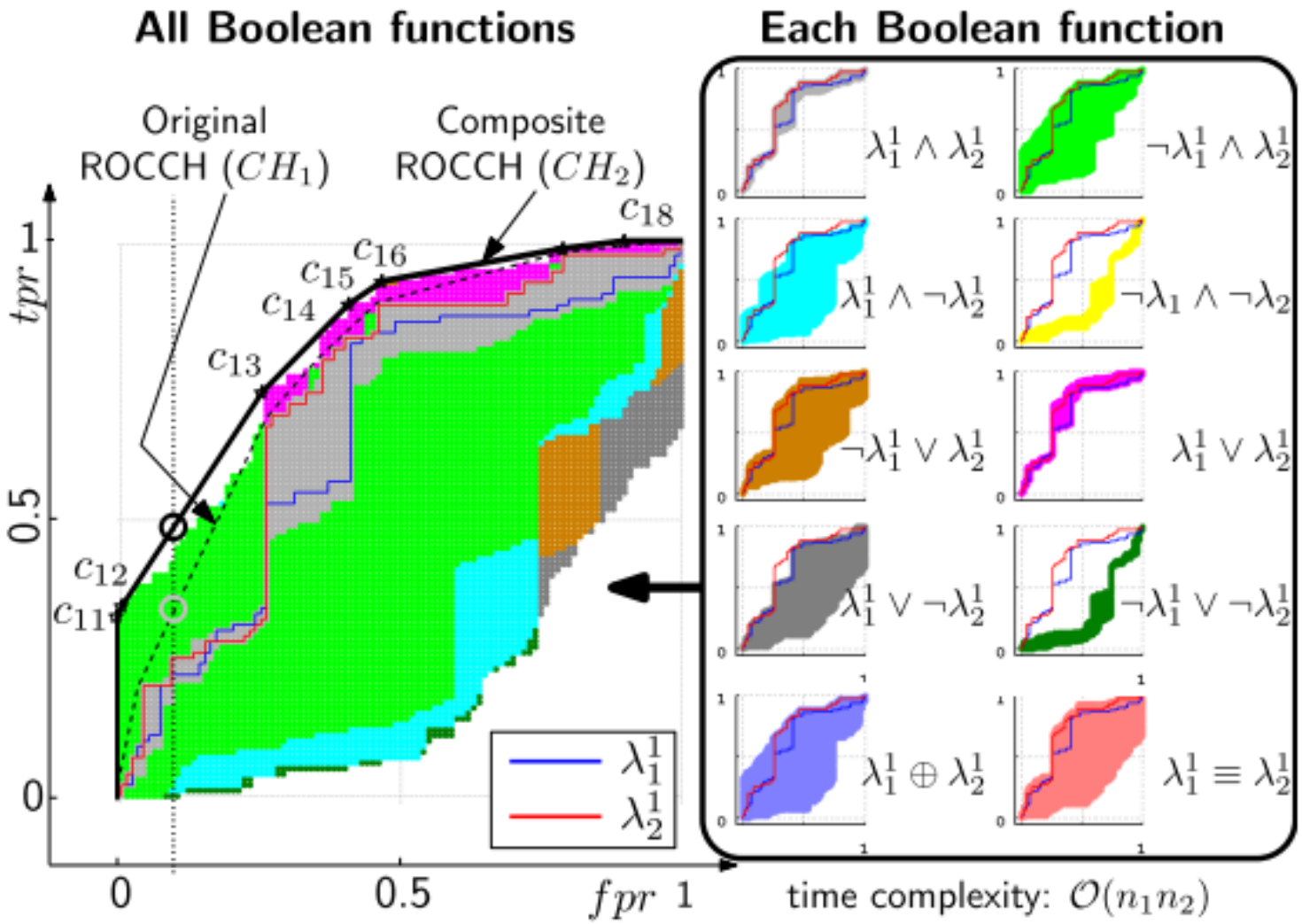
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

IBC: Combinaison booléenne itérative dans l'espace ROC

- Pour chaque seuil du premier détecteur et chaque seuil du deuxième détecteur:
  - Combiner les réponses en utilisant toutes les fonctions booléennes
  - Sélectionner des seuils et des fonctions booléennes tout en améliorant l'espace ROC

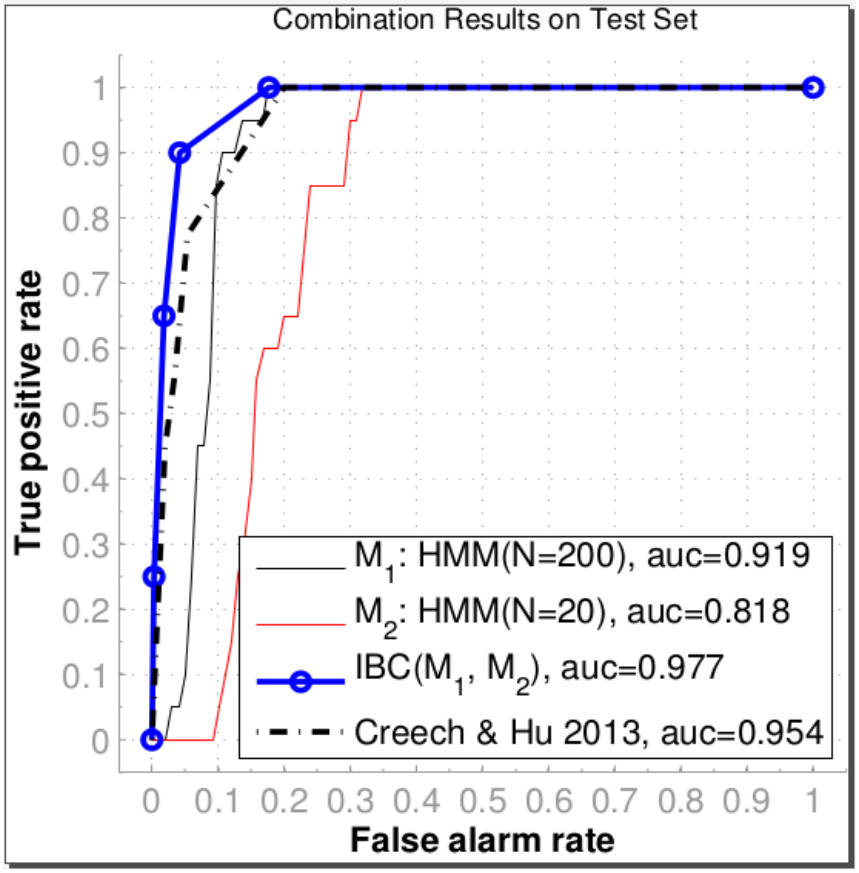
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

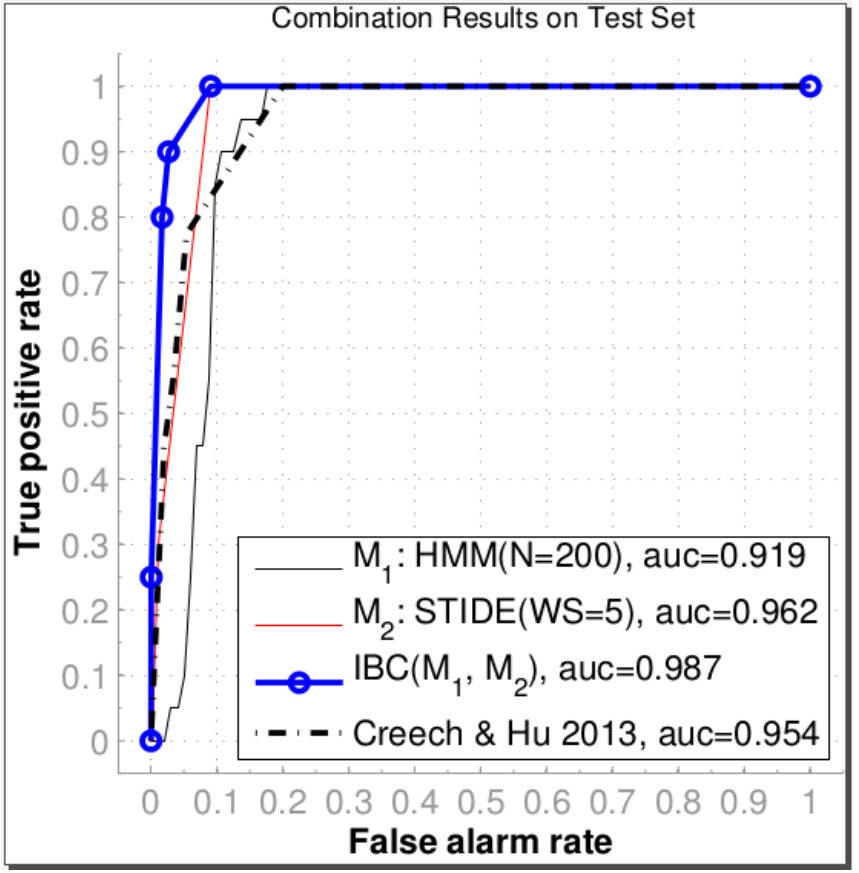
IBC - Exemple



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

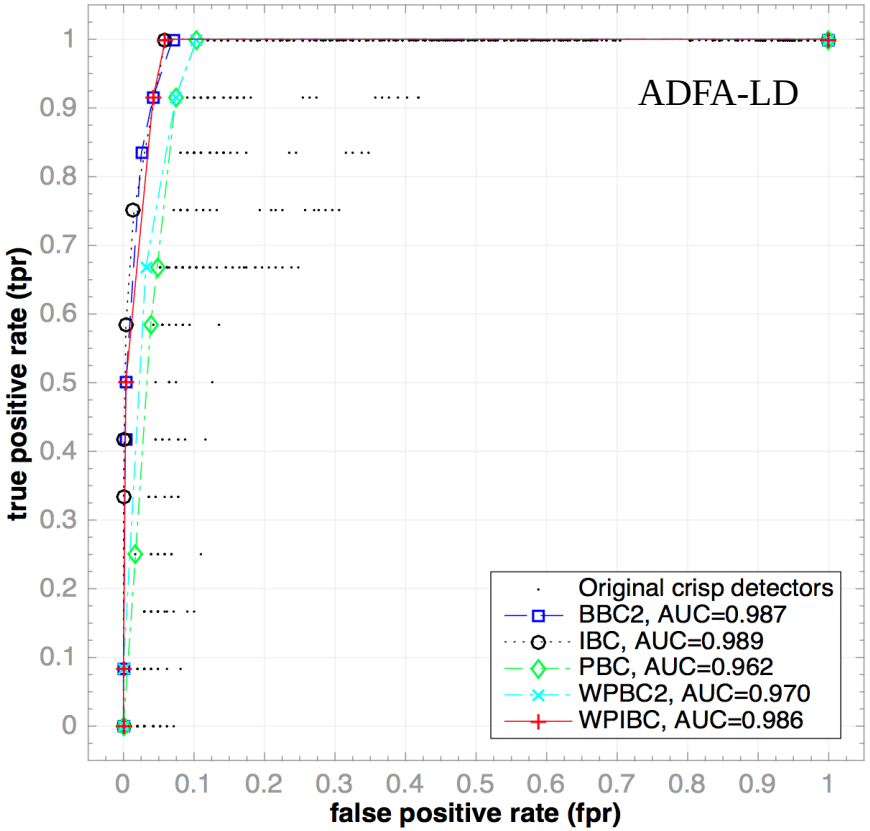
Résultats sur ADFA-LD

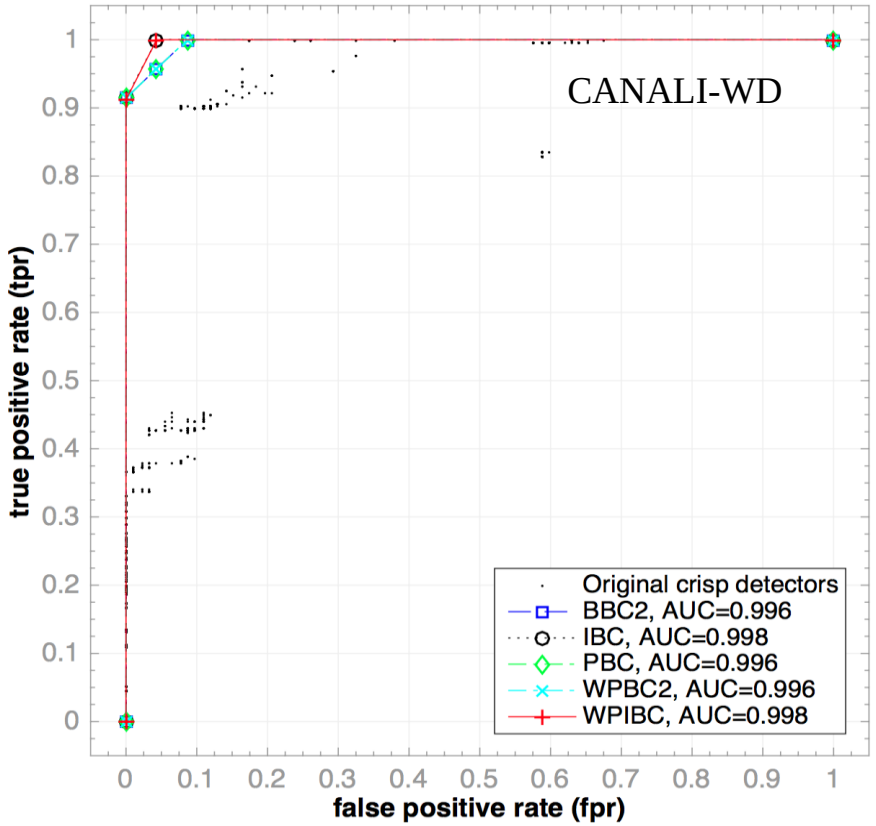




Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Amélioration en utilisant les techniques de filtrage





Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

An Anomaly Detection System based on Ensemble of Detectors with Effective Pruning Techniques

Amirreza Soudi, Wael Khreich, and Abdelwahab Hamou-Lhadj
Department of Electrical and Computer Engineering,
Concordia University, Montreal, QC, Canada
Email: {am_soudi, wkhreich, abdelw}@ece.concordia.ca

A Host-based Anomaly Detection Approach by Representing System Calls as States of Kernel Modules

[1]Syed Shariyar Murtaza, [1]Wael Khreich, [1]Abdelwahab Hamou-Lhadj, [2]Mario Couture

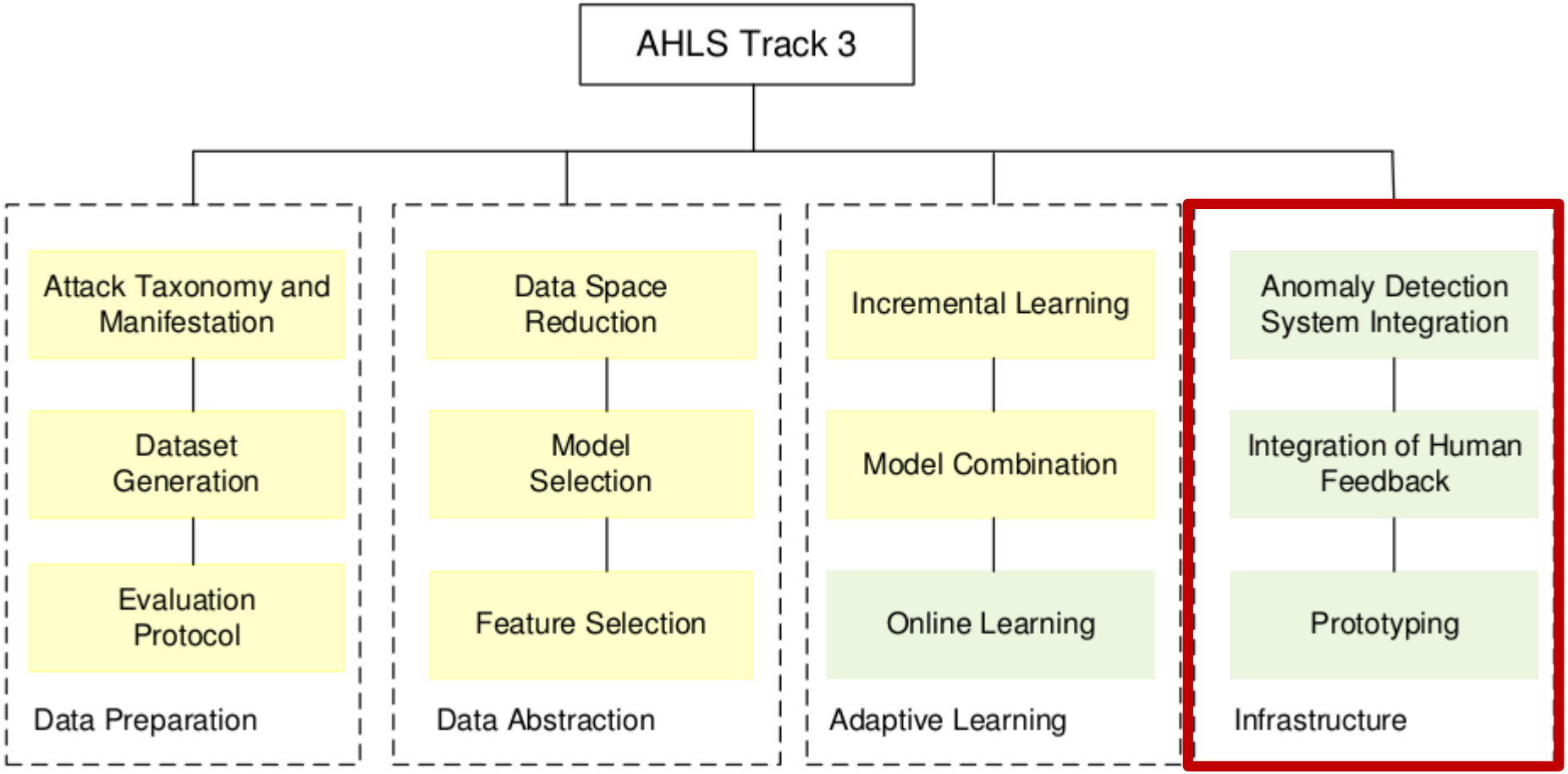
Anomaly Detection Techniques Based on Kappa-Pruned Ensembles

Md. Shariful Islam, Wael Khreich, and Abdelwahab Hamou-Lhadj, *Member; IEEE*

Combining Heterogeneous Anomaly Detectors for Improved Software Security

Wael Khreich[a,*], Syed Shariyar Murtaza[a], Abdelwahab Hamou-Lhadj[a], Chamseddine Talhi[b]

Structure du projet



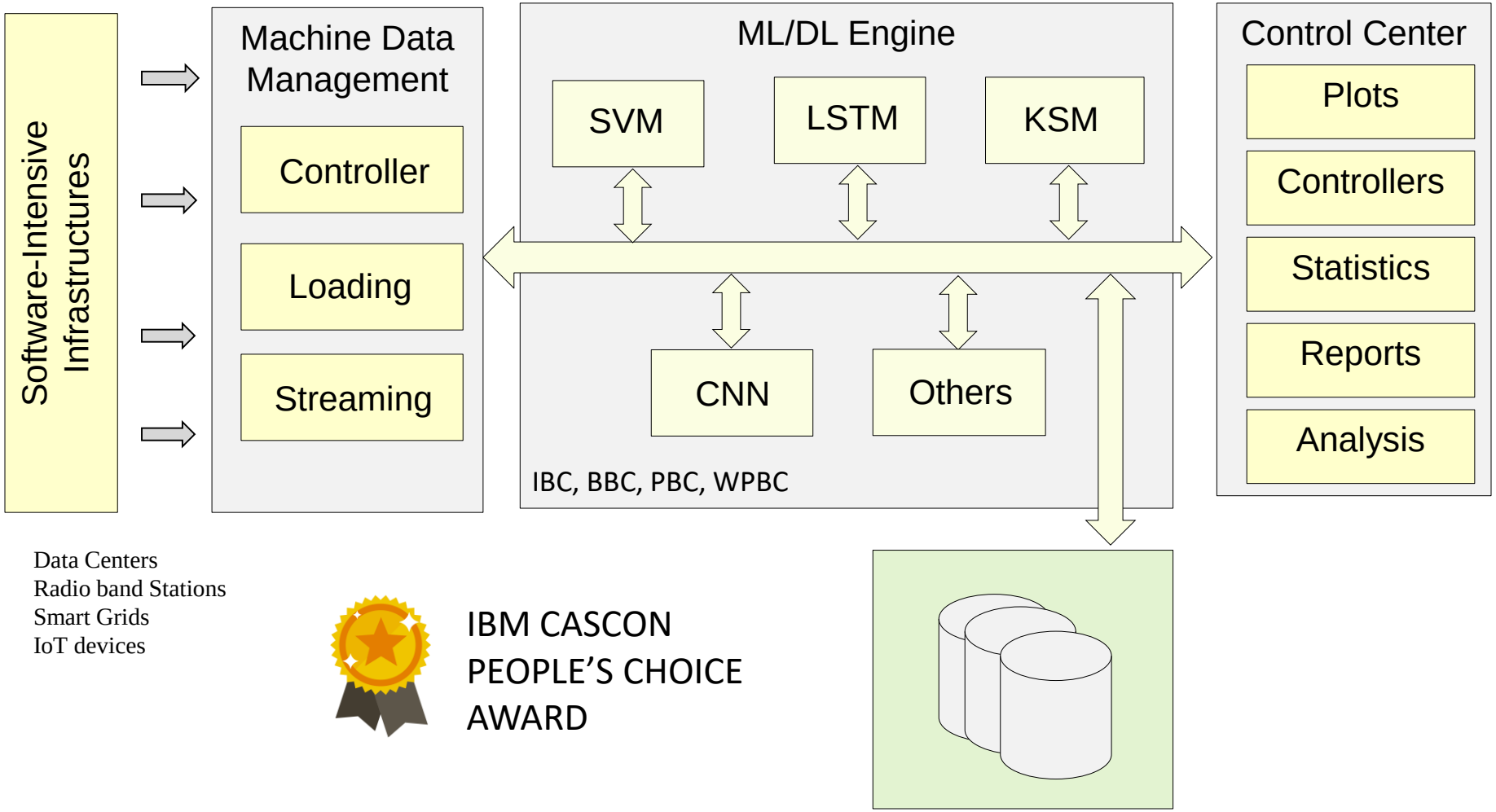
Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Système de détection d'anomalies TotalADS

- TotalADS est une plateforme intégrée pour la détection d'anomalies
  - Plug-in Eclipse
  - Basé sur TMF (TraceCompass)
  - Variantes de STIDE, HMM, KSM, SVM, etc.
  - Analyse de traces
  - Analyse des causes
  - Différents formats de traces
  - Techniques de visualisation

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

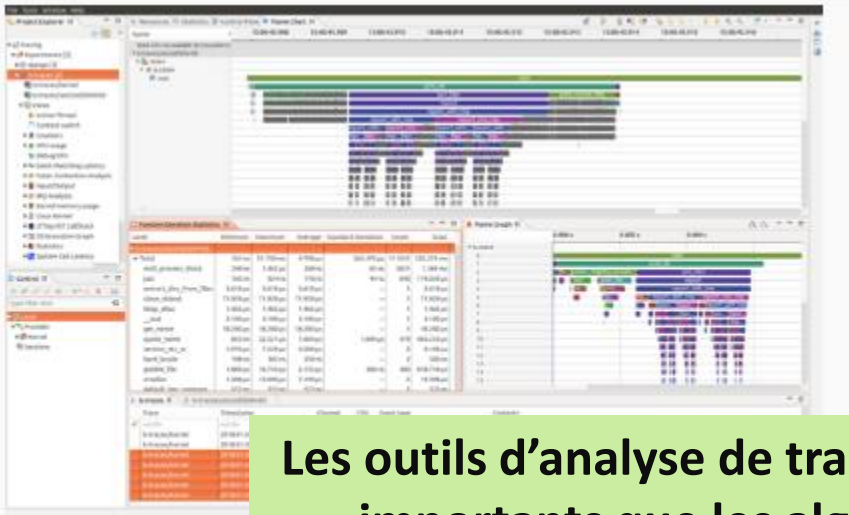
Architecture de TotalADS

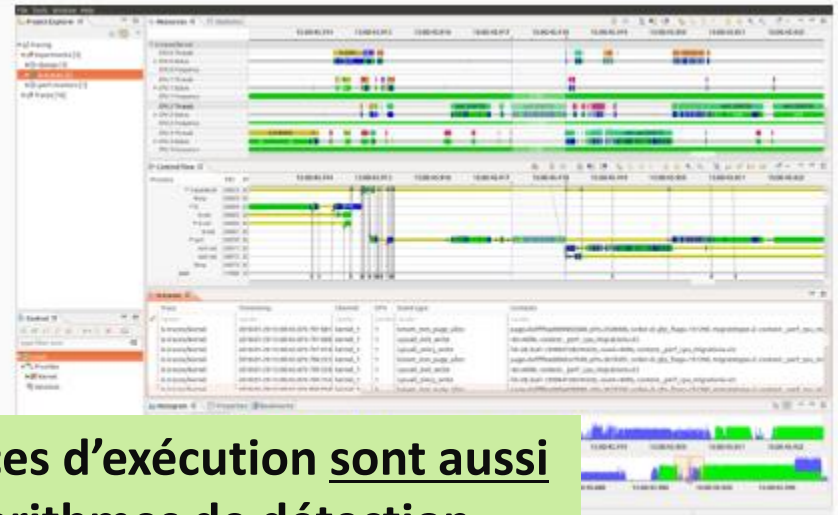


Data Centers
Radio band Stations
Smart Grids
IoT devices

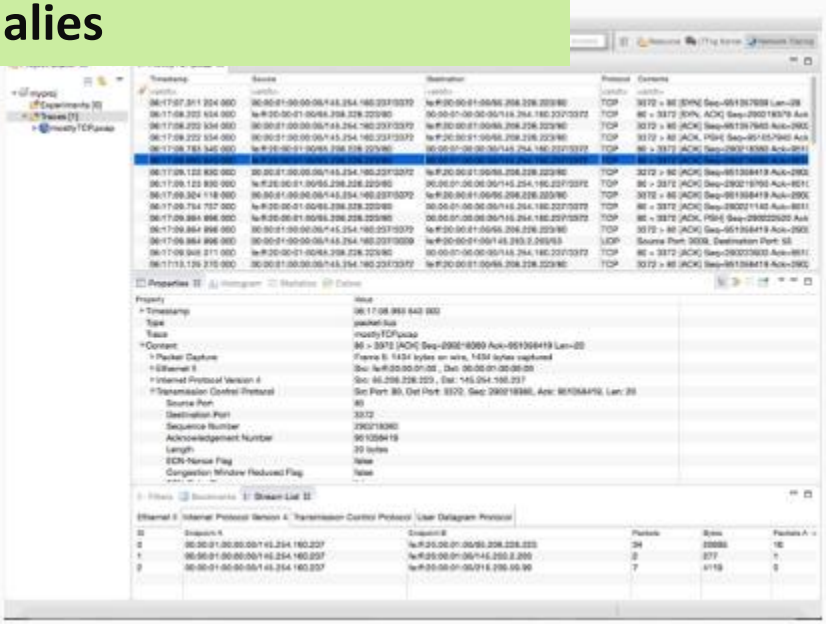
IBM CASCON
PEOPLE'S CHOICE
AWARD

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)





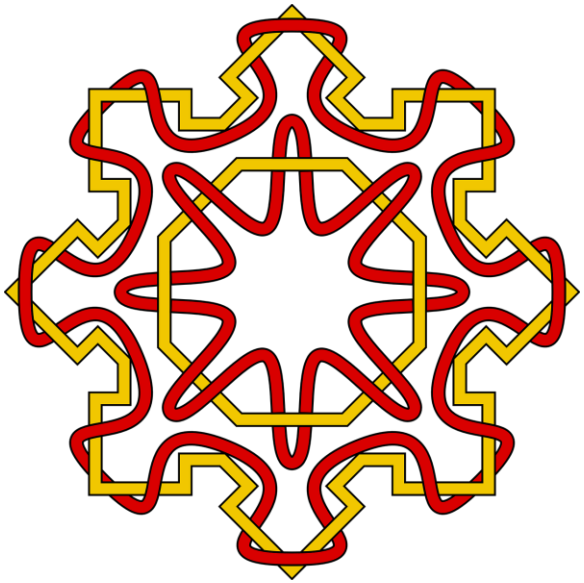
Les outils d'analyse de traces d'exécution <u>sont aussi importants</u> que les algorithmes de détection d'anomalies



Défis, limites et quelques solutions pratiques

Complexité d'entrainer et de tester les programmes DL

- Sélection de modèle/fonctionnalité n'est pas évidente
- Procédure d'optimisation des paramètres
- Ajustement excessif des modèles
- Tester les programmes DL est un problème complexe



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

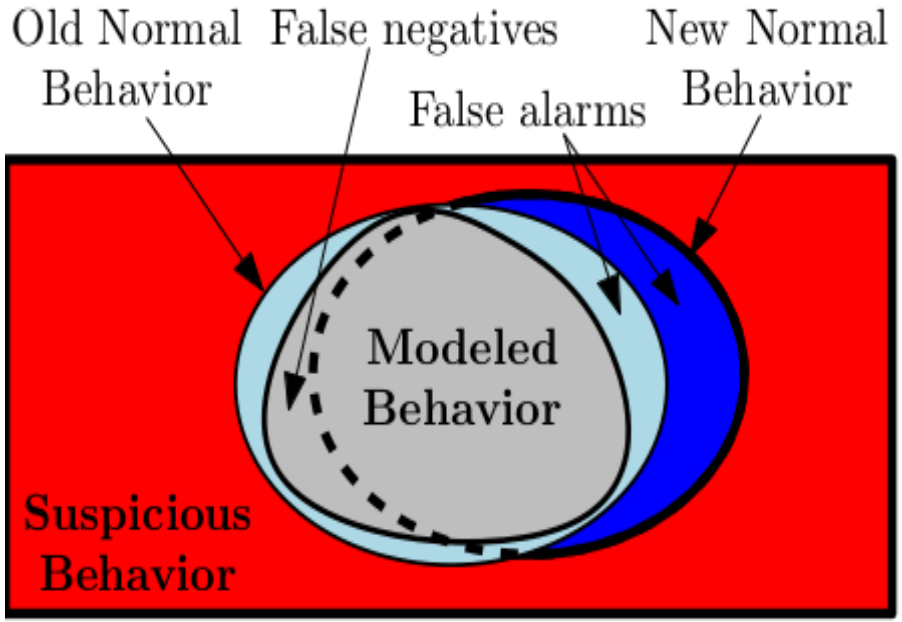
Les jeux de données d'évaluation

- Pas assez de données
- Processus de génération des données rarement documenté
- Les appels systèmes offrent une vue limitée du système
- Manque de protocole d'évaluation

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Mise à jour des systèmes

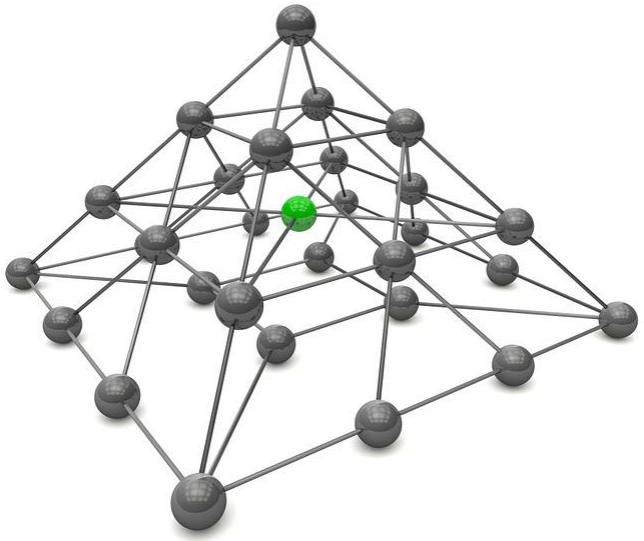
L'intégration continue et mises à jour fréquentes des systèmes informatiques exigent à ce qu'on repense complètement la conception des HADS



Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Évolution de l'environnement hôte

- Les applications ne sont plus monolithiques
- Architecture micro-services
- La virtualisation
- L'utilisation des contenants
- Systèmes distribués et parallèles



Les technologies émergentes

- Internet des objets
- Réseaux électriques intelligents
- Véhicules autonomes
- Systèmes cyber-physiques
- Technologies mobiles

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Quelques solutions pratiques

- Concentrer les travaux de recherche sur des scénarios pratiques en favoriser la collaboration industrie/gouvernement/université
- Investir dans la collecte d'une quantité importante de données
- Développer des protocoles d'évaluation des techniques ADS
- Les outils DL doivent être observables (au cas de problèmes)
- Intégrer les techniques ADS avec les autres mécanismes de sécurité
- Favoriser le développement et partage des outils open source
- Former des équipes multidisciplinaires de développeurs, scientifiques de données, experts des domaines, etc.

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

Conclusion

- Les études récentes montrent que les techniques DL sont prometteuses pour les HADS
- Des études supplémentaires et plus approfondies, avec plus et meilleures données, sont nécessaires pour bien comprendre les gains
- La recherche dans ce domaine doit relever le défis liés à la complexité des systèmes, les mises à jour fréquentes, l'émergence des nouveaux paradigmes, etc.
- Tester et debugger les outils DL sont des activités aussi importantes que le développement de ces outils

Dr. Wahab Hamou-Lhadj (wahab.hamou-lhadj@concordia.ca)

CONCORDIA.CA

Evaluation: Australian Defence Force Academy (ADFA)[1] Linux Dataset

- Ubuntu 11.04, Apache 2.2.17, PHP 5.3.5, TikiWiki 8.1, FTP server, MySQL 14.14 and an SSH server
- Web-based exploitation
- Simulated social engineering
- Poisoned executable
- Remotely triggered vulnerabilities
- Remote password brute force attacks
- System manipulation
- Etc.

[1]http://www.cybersecurity.unsw.adfa.edu.au/ADFA IDS Datasets/