

SECURITE ET ANALYSE DE TRAJECTOIRES AEROPORTUAIRES



Celia Kherfallah Stieban Fernandez

Encadrants : Nicolas Baskiotis, Vincent Guigue

PLDAC 2017-2018

Table des matières

Introduction	2
Prétraitement des données	3
Indexation et caractérisation de l'espace.....	4
Indexation	4
Caractérisation	4
Modélisation, extraction et description des situations.....	6
Modélisation d'une situation	6
Extraction de situations	6
Description et Implémentation des features.....	7
Analyse des situations	9
T-SNE	9
Isolation Forest.....	10
Toolbox	10
Conclusion.....	12
Annexes.....	13
Bibliographie.....	16

Introduction

La sécurité aérienne reste jusqu'à présent la priorité numéro un de l'industrie aéronautique. Des avancées en matière de technologie de navigation aérienne ont réduit de manière significative les collisions en vol et les crashes sur les montagnes. Parmi celles-ci, citons la technologie TCAS [1] - Traffic Collision Avoidance System - qui a permis de réduire le risque de collision en vol. Elle est aujourd'hui exploitée sur tous les avions de transport.

Les accidents d'avions ont longtemps été associés à des incidents en vol. Néanmoins, les collisions au sol entre avions ou impliquant un avion et un véhicule roulant dans l'aéroport existent et représentent un réel danger. Par ailleurs, rappelons l'accident le plus meurtrier de l'histoire de l'aviation civil. En 1977. Une collision au sol à l'aéroport de Tenerife aux Îles Canaries entre deux Boeing 747, l'un de la PanAm et l'autre de KLM, cause la mort de 583 personnes. Cette catastrophe a démontré que la sécurité ne se limite pas que dans les airs mais également au sol.

Par conséquent, les mesures de sécurité et de prévention ont été renforcées. Des systèmes plus rigoureux ont été développés tels que le système de balisage nocturne basé sur la technologie ARCAL [2] - Aircraft radio control of aérodrome lighting - Un système de conseil automatique conçu pour prévenir les accidents et réduire le nombre d'incursions sur piste tout sans interférer avec les opérations aéroportuaires. Il est conçu pour être compatible avec les procédures existantes et comprend des feux d'entrée de piste (RELS-Runway Entrance Lights) et des feux de décollage (THLs- Take off Hold Lights)

Des heuristiques d'espacement d'avion sur les pistes selon leurs poids et leurs catégories restent encore un sujet ouvert aux chercheurs. En effet, les appareils gros porteurs génèrent de fortes turbulences dans leurs sillages [3] (turbulences aérodynamiques qui se forment à l'arrière d'un aéronef) qui peuvent être dangereuses pour les avions suiveurs. La traînée d'air chaude (moins dense que l'air ambiant) peut provoquer la sortie d'un avion plus léger. Dans les cas extrêmes, ce dernier peut être renversé ou recevoir des débris dangereux (FOD - Foreign Object Damage) soulevés par la puissance du souffle du premier avion. Finalement, une distance de sécurité plus importante devra être respectée entre les avions lors du décollage et la phase d'approche, d'où la nécessité de trouver un bon scénario afin de minimiser le retard tout en respectant certaines contraintes.

Certes, ces procédures ont permis un meilleur contrôle du trafic au sol, ayant réduit le nombre de véhicules ou avions s'engageant par erreur sur une piste d'aéroport. Néanmoins, l'expansion du trafic aérien entraînera dans un futur proche une importante saturation au sol. Ainsi, les heuristiques actuelles utilisées par les contrôleurs seront insuffisantes pour gérer le surplus d'activité. C'est pourquoi, la sécurité des pistes est devenue un domaine d'intérêt majeur pour l'industrie, en raison de la fréquence des incidents et/ou accidents dans les environnements de congestion. Il s'agit notamment de collisions, de dépassements et d'incursions sur piste ou voie de circulation. Cette dernière est la catégorie d'accidents la plus fréquente (soit 22% de tous les accidents au cours de la période 2010-2014 [4]).

Il est donc nécessaire de développer des systèmes intelligents, capables de détecter les anomalies qui seront traitées par des experts en aéronautique, en vue de les corriger et améliorer la sécurité.

Notre étude est quant à elle portée sur les points suivants :

- Prétraitement des données.
- L'indexation spatiale de l'aéroport.
- Modélisation, description et extraction des situations.
- Conception d'une interface logicielle afin de visualiser des données.
- Isolation des anomalies.

Dans un premier temps, on modélise une situation par deux trajectoires d'avion au sol, s'étant produites dans la même fenêtre temporelle et spatiale.

On définit aussi une anomalie comme étant une situation ou la fin du scénario qui peut aboutir à un incident ou à un accident (incursions sur piste ou voie de circulation, collisions...).

Nous mettons en place un logiciel offrant plusieurs fonctionnalités :

- Extraction intelligente des situations. Celle-ci est paramétrable afin que les experts du domaine puissent définir à leur guise les critères spatio-temporels.
- Modélisation d'une situation par un ensemble de descripteurs qui peuvent être définis par l'utilisateur.
- Visualisation et analyse des situations par le biais d'une interface graphique permettant la détection de situations anormales.

Dans le cadre de ce projet, les données sont principalement issues des trajectoires d'avions du premier et quatrième trimestre de l'aéroport Charles de Gaulle (Paris) durant les années 2014-2015.

Les principaux axes de notre étude sont :

1. Caractérisation de l'aéroport

Nous procédons à une indexation de l'aéroport dans l'objectif d'attribuer une sémantique aux différentes zones de l'aéroport. Une présentation discrète nous permettra d'obtenir un découpage de l'aéroport en sous régions (terminal, taxiways, voie de traverse et piste de décollage/atterrissage).

2. Modélisation et extraction des situations

Suite au découpage des zones de l'aéroport effectuée dans l'axe 1, l'extraction des situations se fera grâce à un système de requête intelligent. De ce fait, nous serons capables d'extraire des situations selon des zones et à des dates spécifiques. L'ensemble des situations extraites représentera par la suite notre base d'exemples.

3. Description des situations

Après extraction il sera question de décrire une situation dynamique (un appareil en mouvement selon une période fluctuante) en une situation statique. Autrement dit de figer chaque infime instant t (point) de cette situation. Néanmoins, la prédilection des fonctionnalités constitue une étape délicate qui requiert des connaissances métiers.

4. Détection de situations à risque (anomalies)

Nous utilisons des algorithmes de Machine Learning afin de détecter les situations dites à risque. Ces dernières étant peu nombreuses et difficiles à détecter du reste, nous proposons de les isoler grâce à deux algorithmes.

- L'algorithme t-SNE (t-distributed stochastic neighbor embedding) : nous permettra une meilleure visualisation des données issues de la base d'exemples, et le regroupement des situations similaires.
- L'algorithme d'Isolation Forest [4] basé sur les arbres de décision pour l'isolation des anomalies dans la base.

Enfin, ce projet aura pour but de contribuer à l'amélioration de la sécurité aéroportuaires à l'aide des outils de la Data Science. Les résultats obtenus seront alors interprétables et pourront être traités par des experts dans la perspective de mise en œuvre de solutions.

Prétraitement des données

Les informations caractérisant une trajectoire sont collectées en moyenne toutes les dizaines de secondes à l'aide d'un dispositif électronique (transpondeur) embarqué dans chaque appareil. Celui-ci nous communique :

- Les coordonnées géographiques de l'avion (GPS).
- La vitesse au sol de l'avion exprimée en m/s.
- L'accélération et orientation du véhicule par rapport au Nord.
- Procédure de tractage.
- Positionnement de l'avion par rapport au rectangle de piste

D'autres informations nous ont été également communiquées telles que les conditions climatiques et le type d'avion.

De part ces données, nous modélisant une trajectoire T_k par un ensemble de point en fonction du temps, chaque point possède les informations citées précédemment.

φ : ensemble d'informations citées précédemment.

α : ensemble d'aéronefs constituant notre base.

$T_k^v = \{ (t_1, l_1), \dots, (t_{|q|}, l_{|q|}) \}$, avec $(t \in \mathbb{R}, l \in \varphi, v \in \alpha)$.

Les données sur lesquelles notre projet s'appuie ont été préalablement lissées par des experts du domaine. Pourtant, nous constatons quelques comportements anormaux, tels que la présence de trajectoires d'une durée de plusieurs jours ou moins de quelques minutes, qui ont peut-être été causées par un dysfonctionnement du transpondeur.

Ainsi, les trajectoires excédant une heure ou durant moins de quelques minutes sont jugées non fiables. Par conséquent, elles ont été retirées.

De plus, des informations sont manquantes, comme la météo nécessaire à la visibilité et le terminal d'où débute la trajectoire.

Ces attributs serviront par la suite à l'implémentation des fonctionnalités. Par conséquent, nous leur attribuons pour la météo (resp. Terminal) la valeur de la trajectoire la plus proche temporellement (resp. Spatialement).

Indexation et caractérisation de l'espace

Indexation

Précédemment, nous avons exprimé une trajectoire comme étant un ensemble de points contigus en fonction du temps. Or en pratique, ce formalisme est peu exploitable pour modéliser les déplacements d'un véhicule. Les trajectoires possèdent un comportement imprévisible, car elles peuvent s'arrêter dans une zone pendant un certain temps, et ainsi avoir plus de points dans cette région. Néanmoins, tous ces points peuvent être assimilés à un seul point, vu qu'ils possèdent les mêmes caractéristiques. Pour pallier à ce problème, nous proposons une discrétisation de l'espace afin de faciliter l'analyse. Une trajectoire peut être alors modélisée par un vecteur de cases indiquant la présence de l'avion dans la case d'indice (i,j) .

D'une part, cette approche permet de regrouper et moyenner les informations de la trajectoire sur une zone géographique particulière. Les données n'étant pas très bien échantillonnées, cette solution permet la comparaison des trajectoires bien que le nombre de points pour chacune d'elles soit différent sur cette zone.

Le choix du pas de discrétisation se fait selon plusieurs contraintes:

1. Les cases doivent assez grandes pour contenir un avion, celui-ci pouvant atteindre jusqu'à 100 m de longueur et 80 m d'envergure.
2. Les cases doivent d'une part suffisamment petites afin de minimiser la perte d'information sur les trajectoires mais d'une autre part assez larges pour éviter la discontinuité des trajectoires mal échantillonnées.

Nous choisissons alors une discrétisation 110 x 57, qui donne une dimension de 100 x 100 m² pour chaque case. Notons que cette dernière est approximative car nous utilisons ici un espace euclidien tandis que l'espace réel est non-euclidien.

Caractérisation

L'indexation permet également la caractérisation de l'espace. Par conséquent, nous avons mis en place une base de données où chaque trajectoire est représentée par un ensemble de cases où sa présence est vérifiée. Ce système permet alors de récupérer par requêtage les caractéristiques d'une ou plusieurs trajectoires dans une ou plusieurs cases.

La moyenne de la vitesse dans chaque case permet d'avoir une première représentation de l'aéroport.

Comme nous pouvons le constater dans la figure ci-dessous, les zones où la vitesse est faible représentent les phases

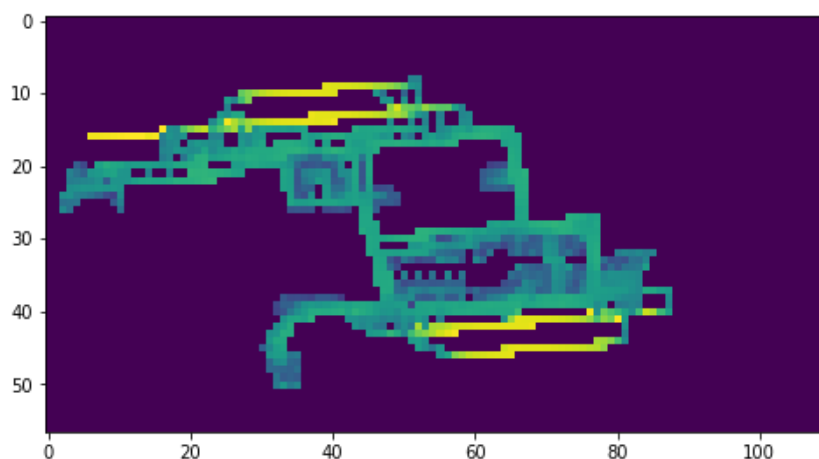


Figure 1 - Carte de l'aéroport avec la vitesse moyenne par case

d'approche, les entrées sur pistes, les sorties de terminaux et hangars

La vitesse étant quantitative, la transformation de cette caractéristique en variable qualitative offre une meilleure distinction entre les cases. Pour ce faire, nous utilisons l'algorithme K-moyennes sur la vitesse moyenne de chaque case en vue de les labelliser.

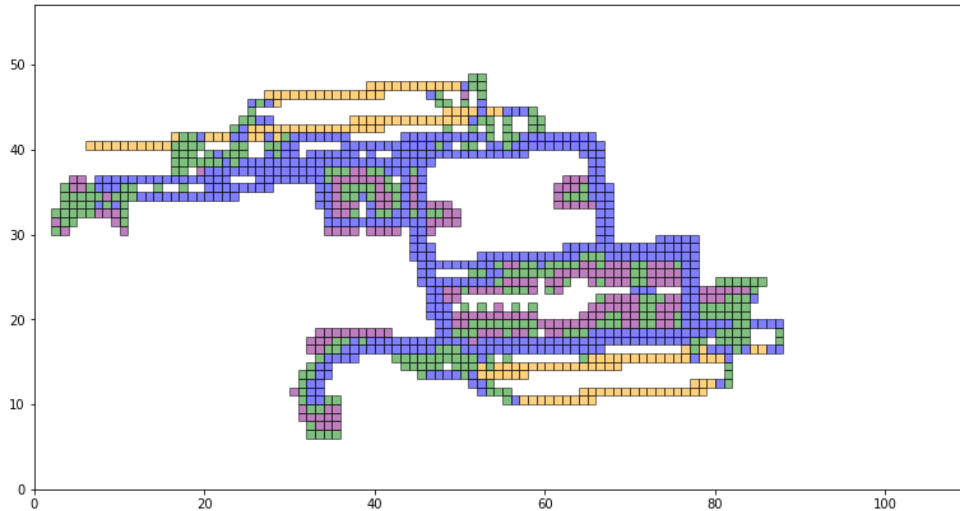


Figure 2 - Labellisation des cases selon la vitesse avec l'algorithme K-means avec K=4

Nous faisons intervenir une autre mesure : la vitesse angulaire.

$$V\theta = \frac{\theta_2 - \theta_1}{t_2 - t_1} [\text{rad/s}]$$

Une valeur élevée (resp. faible) détermine une trajectoire courbée (resp. droite). Cette nouvelle information nous permet de déceler les intersections qui sont des zones pertinentes. Le risque collision est particulièrement élevé dans ces dernières.

2 CARTE CASES VITESSE ANGULAIRE K-MEANS (k=2)

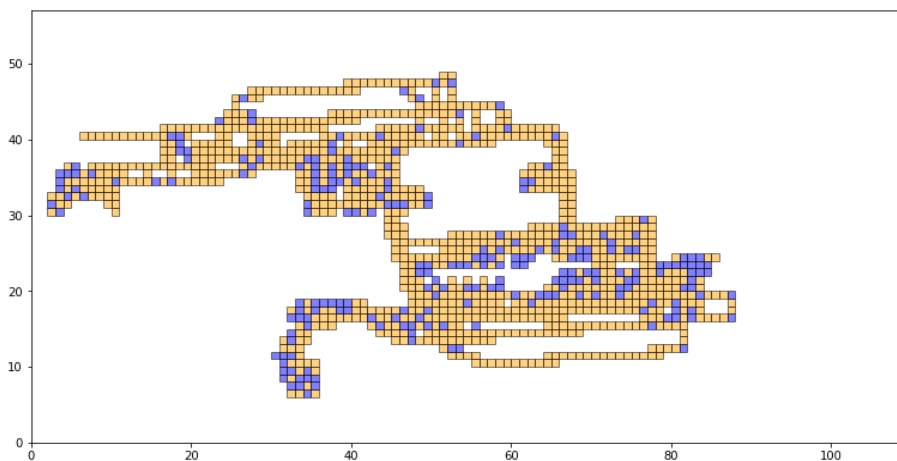


Figure 3 - Application de K-means sur la vitesse angulaire avec K=2

Modélisation, extraction et description des situations

Modélisation d'une situation

Nous définissons une situation comme étant un couple de trajectoires partielles associées à une paire d'avions au sol, qui se sont produites dans la même fenêtre temporelle et spatiale, En d'autres termes, le moment auquel les deux avions sont relativement proche l'un de l'autre.

Soit $S_{v,w} = \{ \{ T_{n:m}^v, T_{p:q}^w \} / n,m,p,q \in \mathbb{N}, v,w \in \alpha, v \neq w \}$, avec $T_{n:m}^v, T_{p:q}^w$ des trajectoires partielles.

Deux avions peuvent se croiser plusieurs fois lors de leurs déplacements, nous distinguons alors ces rencontres et les considérons comme étant indépendantes.

Soit deux aéronefs a_1, a_2 s'étant rencontré à des dates t_1 et t_2 à deux emplacements d_1, d_2 avec $t_1 - t_2 < \text{temps_minimal [s]}$ et $d_1 - d_2 < \text{distance_minimale [m]}$.

Ainsi les portions de trajectoires j_1 et j_2 associées aux deux avions sont des candidats potentiels pour modéliser une situation. Néanmoins, deux formulations de la situation sont possibles : $S_{12} = (j_1, j_2)$ et $S_{21} = (j_2, j_1)$

Pour éviter toute ambiguïté et risque de fausser les résultats de nos modèles, nous imposons une symétrie. Par conséquent, la description de la situation se doit d'être relative aux deux véhicules, les situations (j_1, j_2) et (j_2, j_1) se doivent d'être identiques.

Une base de situations est constituée de plusieurs situations décrites homogènement.

$\beta = \{ S_{v,w} / v,w \in \alpha, v \neq w \}$.

Extraction de situations

L'extraction de situations passe par des critères spatio-temporels. Pour la condition spatiale, nous déterminons une distance maximale séparant deux aéronefs, en dessous de laquelle, on peut considérer une situation. Cette distance peut être déterminée par deux mesures : la distance euclidienne ou la distance réelle.

Étant donné notre discrétisation de l'espace, pour commencer nous décrivons chaque situation par un ensemble de paires de cases. Chaque paire contient alors les informations relatives à chacune des trajectoires dans leurs cases respectives « au même moment ». Nous proposons pour cela un algorithme combinatoire effectuée en plusieurs étapes.

Etape 1 :

Pour chaque jour, lister les trajectoires qui ont débuté ce jour et la première heure du jour suivant. Former des combinaisons de 2 avec la liste obtenue. L'ajout d'une heure supplémentaire permet de prendre en considération les situations ayant pu avoir lieu autour de minuit.

Etape 2 :

Pour chaque paire de trajectoires (t_1, t_2) , vérifier si la date de départ de t_1 est comprise dans l'intervalle de temps où a eu lieu t_2 . De même pour t_2 dans t_1 . Si ce n'est pas le cas, passer à la paire suivante.

Etape 3 :

Effectuer la liste des cases de t_1 et t_2 dans l'intervalle de temps commun aux deux trajectoires. Soient $|c_1|$ et $|c_2|$ les tailles respectives des cases de t_1 et t_2 , et c_{1i} et c_{2i} les $i^{\text{èmes}}$ cases respectives de t_1 et t_2 .

Algorithme :

```

i, j ← 1, 1
situations ← []
Tant que i < |c1| and j < |c2| :
    d1, d2 ← date de c1i, date de c2j
    dist ← distance(c1i-1, c2j-1)
    Si dist < dist_max alors :
        Ajouter (c1i-1, c2j-1, dist) à situations
    Si d1 < d2 alors :
        i ++
    Sinon :
        j ++
retourner situations

```

Etape 4 :

La dernière étape consiste à considérer chaque rencontre entre deux avions comme une situation unique. La différenciation se fait également par l'intermédiaire de critères spatio-temporels.

Description et Implémentation des features

Le choix des fonctionnalités et leur implémentation constitue une étape délicate qui requiert une bonne connaissance du domaine. Nous proposons une série de features suffisantes à la construction de notre modèle.

1. Distance minimale : Représente la distance minimale en mètre des deux appareils a1, a2 dans la situation $S_{1,2}$ le point associé est appelé le point critique de la situation.

$dist_min = \min \{ d_{i,j} / \forall (i,j) \text{ q } X \text{ q' avec } t_i - t_j < \text{temps minimal} \}.$

2. Pushback : Stipule si les deux avions se font remorqués ou pas. Les pilotes n'ayant aucune visibilité vers l'arrière, des véhicules spéciaux appelés push s'occupe de tracter les avions pour les éloignés des terminaux ou hangars de maintenance leur permettant par la suite de circuler de façon autonome.

3. Types d'avions : Stipule si les deux avions sont de même catégorie, le choix de cette fonctionnalité est dû au problème de l'espacement entre les deux avions, si l'un des deux est plus gros que l'autre une distance de sécurité doit d'être respectée durant le décollage/atterrissage.

4. Area Runway : Stipule si l'avion est dans le triangle de piste.

5. Météo : Reporté les conditions météorologiques d'un jour donné dans la base. En effet, les risques d'accidents sont plus élevés s'il y'a moins de visibilité.

6. Temps total de la rencontre : Représente le temps total écoulé durant la rencontre.

7. Nombre de rencontres : Représente le nombre de fois ou une paire d'avions s'est rencontré durant leurs déplacements au sol.

8. Similarité direction : Stipule si les deux appareils ont la même direction.

9. Angle : Représente l'angle formé par la direction des deux trajectoires.

Soit θ_1, θ_2 les angles par rapport au Nord des deux appareils a1 et a2.

$$\theta_{max} = \max \{ \theta_1, \theta_2 \}, \theta_{min} = \min \{ \theta_1, \theta_2 \}$$

$$\theta_{diff} = \min (\theta_{max} - \theta_{min}, \theta_{min} + 360 - \theta_{max})$$

10. Produit Scalaire et vectorielle des vecteurs de déplacement : Soient les vecteurs de déplacement v1, v2 des appareils a1 et a2, avec comme origine les coordonnées cartésiennes associées respectivement aux points critiques des trajectoires t1 et t2.

L'application d'un produit scalaire sur ces deux vecteurs retourne :

- 1 si v1 colinéaire à v1 et v1 et v2 se dirigent dans le même sens.
- -1 si v1 colinéaire à v1 et v1 et v2 se dirigent dans un sens différent.
- x sinon / $x \in \mathbb{R}$.

L'application d'un produit vectoriel sur ces deux vecteurs retourne :

- 0 si v1 est perpendiculaire à v2.
- x sinon / $x \in \mathbb{R}$.

11. TypeCase : Cette fonctionnalité nous informe sur quel type de case la situation S12 s'est produite. Cette information nous est donnée grâce à la labellisation des cases par l'algorithme k-means en fonction :

- De la vitesse.
- De la vitesse angulaire.

12. Dépassement de vitesse : La discrétisation des cases nous permet de dresser des histogrammes de vitesse (resp. Vitesse angulaire) et ainsi d'estimer si la vitesse de l'appareil est élevée par rapport à la moyenne.

13. DeltaV_par_Vmax : Cette fonctionnalité décrit la variation de la vitesse relative aux deux avions. Soit v1, v2 les vitesses des appareils a1 et a2 aux instant t1 et t2.

$$\text{vitesse_relative} = |V2 - V1| / \max(V2, V1)$$

On constate trois cas:

1. Cas 1: $v1 > v2, v2 = 0$ m/s (et inversement)
vitesse_relative = $|V1| / V1 = 1$

Cette situation stipule que l'appareil a1 est en mouvement et que l'appareil a2 est à l'arrêt.

2. Cas 2: $v1 = v2, v1 \neq 0$
vitesse_relative = $0 / v1 = 0 / v2 = 0$ m/s

Cette situation stipule que les deux appareils sont en mouvement.

3. Cas 3: $v1 = v2, v1 = 0$.
Si vitesse maximale ~ 0 et donc vitesse_relative $\sim \text{inf}$: on remplace cette valeur par -1
Cette situation stipule que les deux appareils sont à l'arrêt.

14. Variation de l'accélération : Cette fonctionnalité nous informe sur la variation de l'accélération des deux véhicules. Soit acc1, acc2 les accélérations des deux l'appareils a1 et a2. La mise en place d'une formule décrivant l'accélération pour les avions nécessite la présence d'un expert. Nous avons choisi de représenter les variances par un vecteur de taille deux.

La 1ere case stipule que le véhicule décélère.

La 2ème case stipule que le véhicule accélère.

Remarque : le cas où l'accélération égale à zéro est omis par cette fonctionnalité car l'information "véhicule à l'arrêt" est déjà spécifiée par le feature Delta_par_Vmax.

15. Vitesse relative : Soit v1, v2 la vitesse des deux appareils au point critique.

$$\text{Vitesse} = \text{sorted}(v1, v2)$$

16. Source terminale : Cette fonctionnalité nous informe de la provenance des deux appareils.

17. Situation_Before / situation_After : Contrairement aux features précédents qui se limitent à décrire la situation à l'instant phare, situation_before (resp. After) calcule la variation de la vitesse, la distance et la durée entre la première (resp. Dernière) case et la case au point critique.

18. Eclairage_du_jour : Stipule si la situation s'est produite la nuit ou le jour.

Analyse des situations

Après la construction de notre base d'exemples ainsi que la mise place des outils de visualisation nous pouvons enfin procéder à la détection des anomalies et les analyser.

Dans le cadre de notre projet, une anomalie peut être vu comme un cas où deux avions au roulage ne respectent pas la distance de sécurité obligatoire, ou une intrusion sur piste alors que l'appareil était en phase de décollage. Ces dernières sont peu fréquentes et différentes du reste du dataset. Afin de les exposer, nous allons appliquer des outils de Machine Learning et les analyser par la suite.

Nous nous basons donc sur les distances minimales obligatoires entre les avions dans l'aéroport, puis effectuons un test à la main.

T-SNE

Nous faisons appel à l'algorithme t-SNE, afin de visualiser graphiquement les données. Plusieurs essais ont été effectués, nous avons varié la perplexité dans le but de mettre en évidence les clusters et testé plusieurs combinaisons de features pour garder les plus pertinents. Si l'absence de celui-ci n'apporte aucun changement, on le considère comme étant non fonctionnel.

Nous faisons appel à l'algorithme t-SNE, afin de visualiser graphiquement les données. Plusieurs essais ont été effectués, nous avons fait varier la perplexité dans le but de mettre en évidence les clusters et testé plusieurs combinaisons de features pour garder les plus pertinents. Si l'absence de celui-ci n'apporte aucun changement, nous le considérons comme étant non fonctionnel.

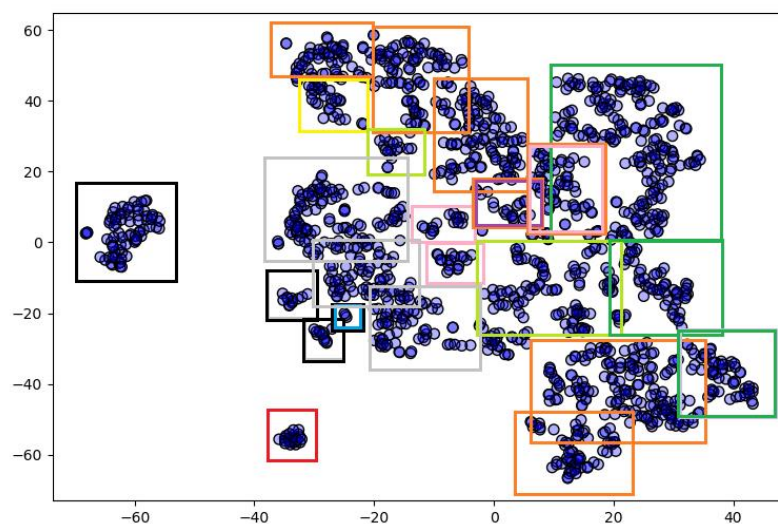


Figure 4 – Affichage du t-SNE avec détection des clusters

Après avoir visualiser manuellement chaque région résultante de l'algorithme T-SNE, nous avons déceler les régions suivantes (Figure 4) :

- **Zone noire** : Représente une situation où l'un des avions sont dans le terminal, un se fait remorquer par le tracteur et l'autre venait tout juste de quitter le parking (gate).
- **Zone verte** : Représente deux avions roulant sur des voies de circulation à une vitesse moyenne comprise entre 7 et 15 m/s.
- **Zone vert-clair** : Ressemble à la zone verte à l'exception que la vitesse moyenne est moins élevée
- **Zone rouge** : Cette région est composée de situations où les avions ont une vitesse tres faible ou une accélération négative. Elles se déroulent principalement autour des terminaux et entrées sur les pistes.
- **Zone jaune** : Cette région comporte les situations se déroule autour des points d'attente et le debut de la piste, la vitesse en moyenne des avions sur la piste dépasse les 20 m/s.
- **Zone orange** : Représente un point d'attente, autrement dit, la zone où les avions attendent d'avoir l'autorisation de la tour de contrôle pour entrer en piste (phase de décollage).
- **Zone bleu-noir** : Est similaire à la zone noire mais aucun avion se fait tracter.
- **Zone rose** : L'ensemble des situations où les avions s'éloignent lentement des terminaux.
- **Zone orange-violet** : Est Similaire à la zone orange, néanmoins l'accélération est négative.
- **Zone grise** : Est un mélange de toutes zones.

Les Figures 9 à 17 fournies en annexes montrent des cas typiques de situations dans chaque zone.

Isolation Forest

Au lieu d'avoir recours à une méthode manuelle, nous avons fait appel à l'algorithme Isolation Forest qui permet de mieux filtrer les situations anormales grâce au score d'anomalie (Figure 4).

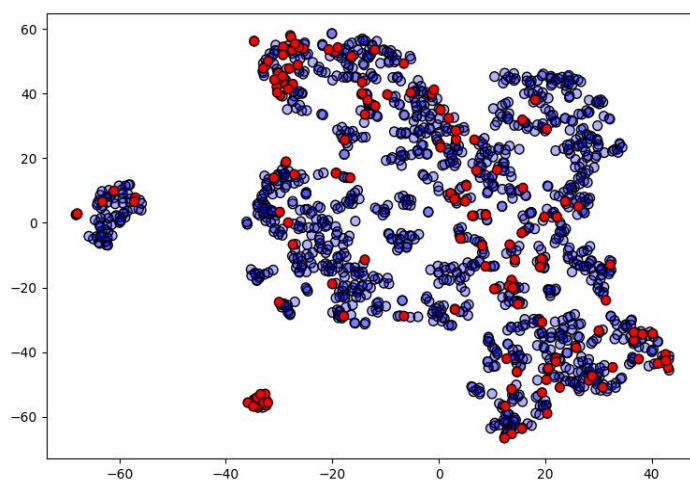


Figure 4 - Résultat de l'Isolation Forest - en rouge les situations anormales.

La situation la plus dangereuse que nous avons pu extraire grâce à l'Isolation Forest s'est déroulée le 16/05/2015. Un Boeing 772 et un Airbus 380 débutent leur roulage vers 21 :40. Tous deux sont de catégorie gros porteur, ils proviennent du terminal 2E et se dirigent vers la piste Ouest de l'aéroport. Pour les situations se déroulant dans les mêmes conditions, nous constatons qu'il y avait toujours un avion qui attendait que la piste se libère. Néanmoins, nous apercevons dans cette situation que les deux avions n'ont pas emprunté la même voie de traverse et que la distance de sécurité obligatoire entre les avions au décollage n'a pas été respectée (un espacement de 2 minutes pour les avions gros porteur). En effet, le B772 est le premier à entrer en piste. Peu de temps après, l'A380 s'engage sur la piste et accélère tellement qu'une distance de

20 mètres les sépare. Par justesse, le A380 ralentit avant d'éviter la catastrophe. Cette situation est une intrusion en piste qui aurait pu causer un accident grave

Toolbox

Différents outils ont été utilisés dans notre étude. Nous avons utilisé MongoDB, un système de gestion de base de données NoSQL, pour le stockage des trajectoires en cases et des situations. Une application logicielle a ensuite été élaborée afin d'effectuer efficacement les analyses.

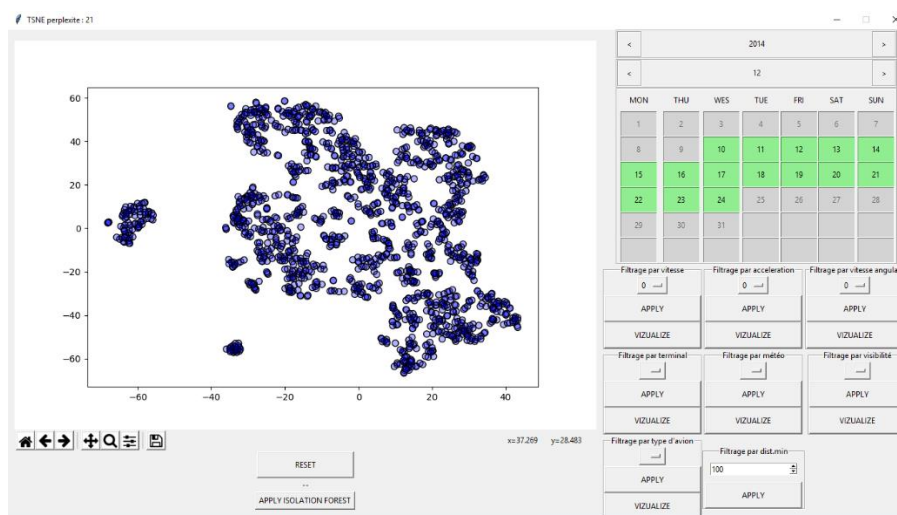


Figure 5 – Fenêtre principale avec la visualisation graphique et les filtres

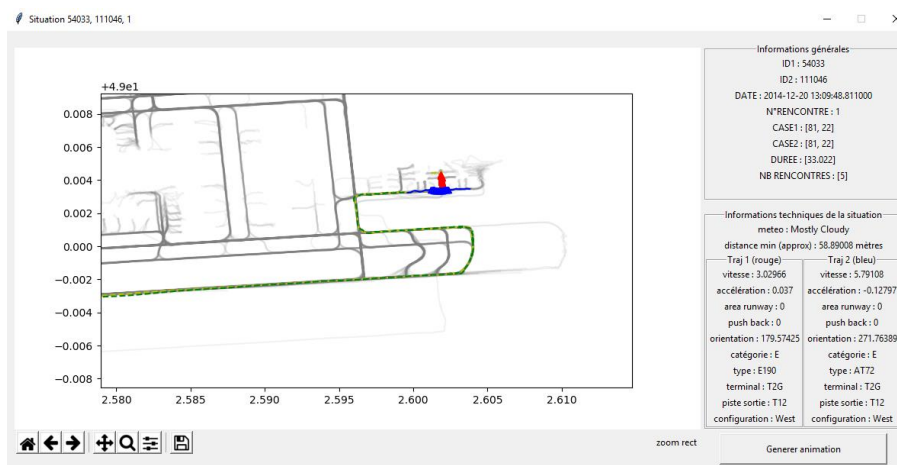


Figure 6 – Fenêtre générée lors du clic sur un point

Une fenêtre graphique offre la visualisation des points générés par le t-SNE. Un calendrier et différents filtres sur les descripteurs des situations sont mis en place pour sélectionner les points et mettre en évidence la répartition des points. Chaque point est cliquable et affiche les informations liées à la situation dans une nouvelle fenêtre. Cette dernière contient une carte de l'aéroport décrivant les deux trajectoires ainsi que leur zone de rencontre. Par ailleurs, il est possible de générer une animation décrivant le parcours des deux trajectoires concernées.

Conclusion

Pour conclure, ce projet avait pour but de détecter les incidents au sol qui représentaient un intérêt majeur pour l'aviation civile. Pour répondre à cela, nous avons proposé de développer un logiciel paramétrable, qui permet de visualiser toutes ces rencontres durant une certaine période de temps, et d'isoler les éventuelles anomalies grâce à des méthodes de Machine Learning. Le prétraitement des données et l'indexation de l'aéroport, étaient une étape importante pour modéliser, décrire et enfin extraire les rencontres des avions au sol. Ce projet nous a permis de sortir du cercle de l'informatique, de découvrir et s'adapter à un autre domaine fascinant qui s'est avéré très constructif.

Annexes

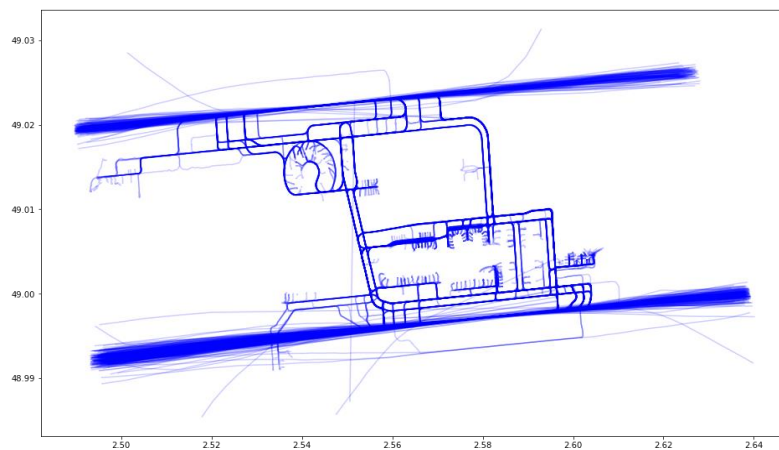


Figure 7 – Carte de l'aéroport générée par les trajectoires

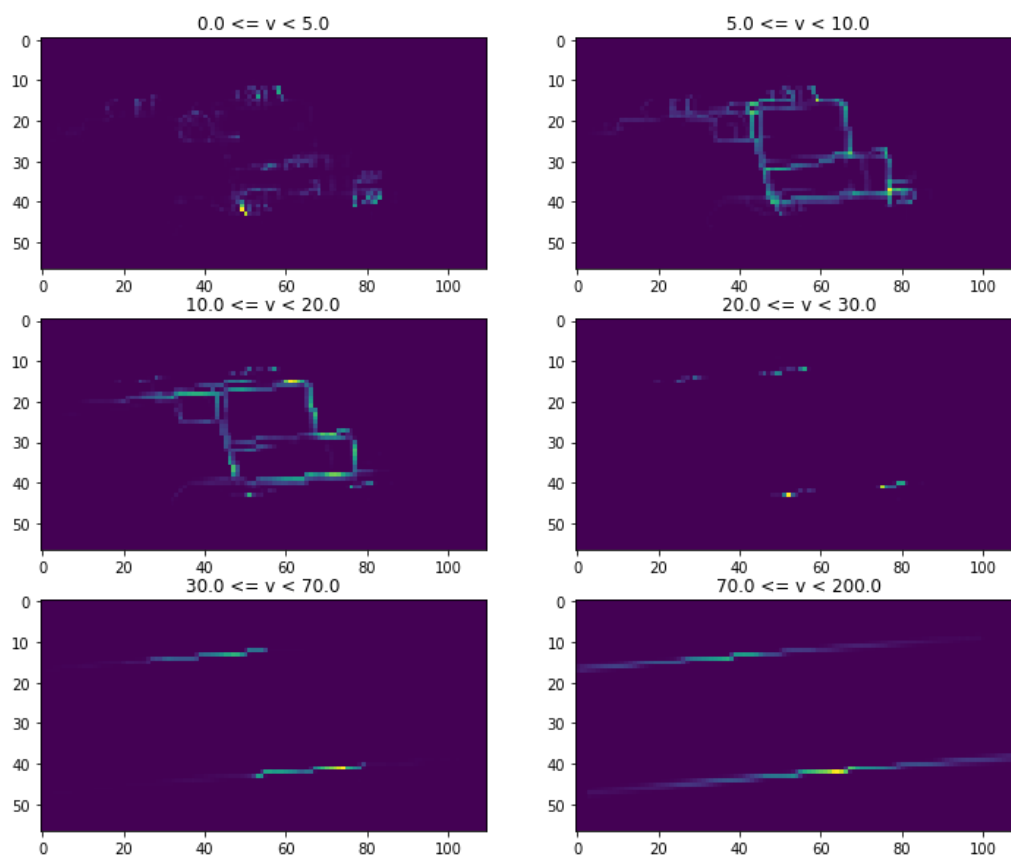


Figure 8 – Distribution de la vitesse discrétisée dans l'aéroport

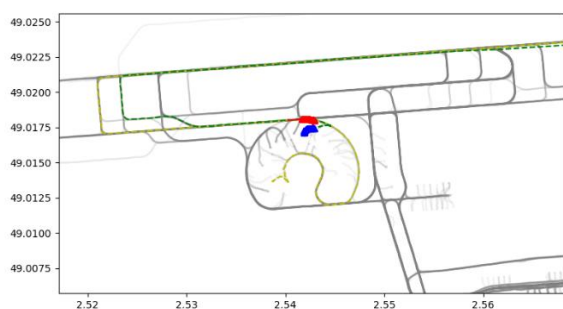


Figure 9 - Zone noire (trajectoire bleue – avion tracté)

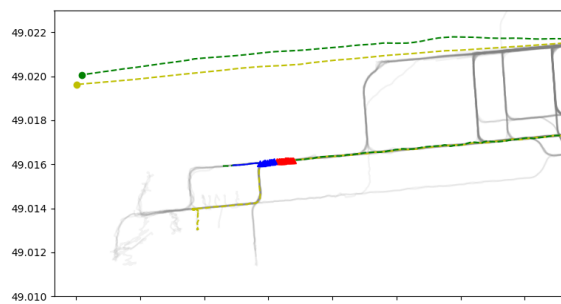


Figure 10 - Zone verte

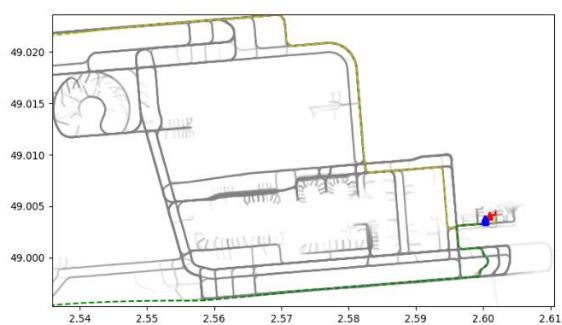


Figure 11 - Zone rouge (deux avions quasiment à l'arrêt)

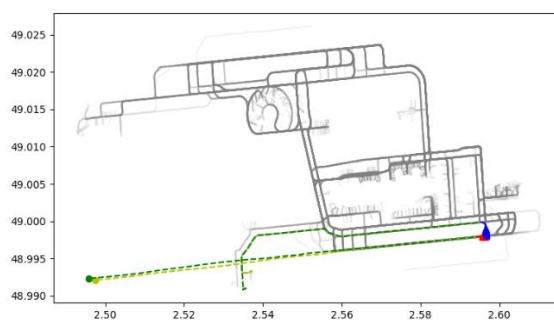


Figure 12 - Zone orange

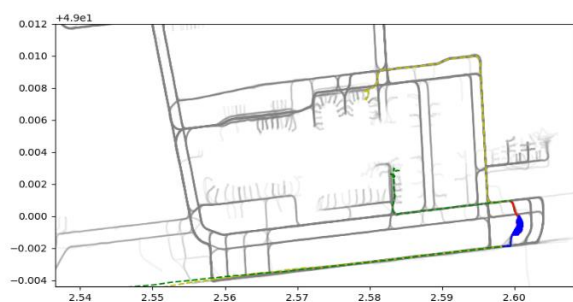


Figure 13 - Zone orange-violet

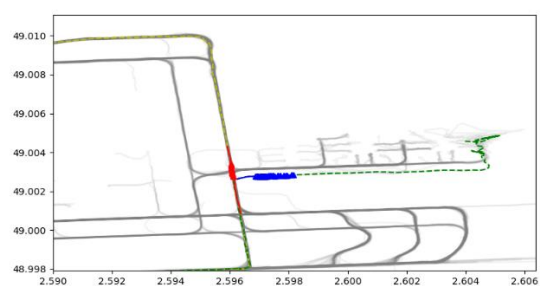


Figure 14 - Zone rose

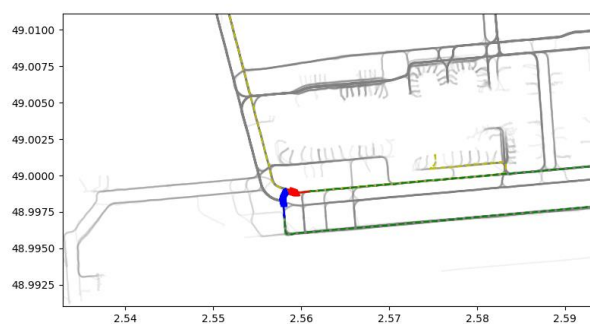


Figure 15 - Zone vert clair

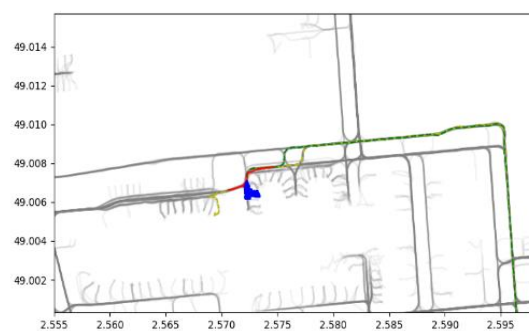


Figure 16 - Zone bleu-noir

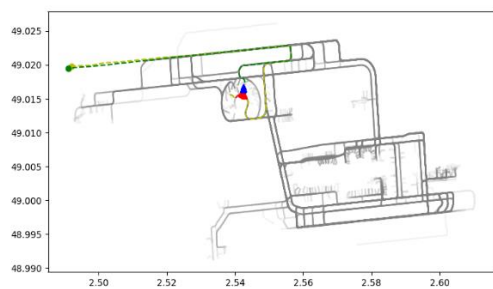


Figure 17 - Zone jaune

Bibliographie :

[1] J. K. Kuchar, « The Traffic Alert and Collision Avoidance System », vol. 16, no 2, p. 20, 2007.

[2] IATA, « Runway Safety Accident Analysis Report », 2015.

[3] J. Fee, « Runway Safety Trends and Runway Incursion Analysis », NTSB Forum on Runway Incursion Safety Issues, September 19-20, 2017.

[4] Z.-H. Zhou, « Fei Tony Liu, Kai Ming Ting Gippsland School of Information Technology Monash University, Victoria, Australia {tony.liu},{kaiming.ting}@infotech.monash.edu.au », p. 10.

[5] F. T. Liu, K. M. Ting, et Z.-H. Zhou, « Isolation-Based Anomaly Detection », ACM Transactions on Knowledge Discovery from Data, vol. 6, no 1, p. 1 39, mars 2012.