Les défis et les opportunités techniques du fonctionnement d'un service antispam mutualisé

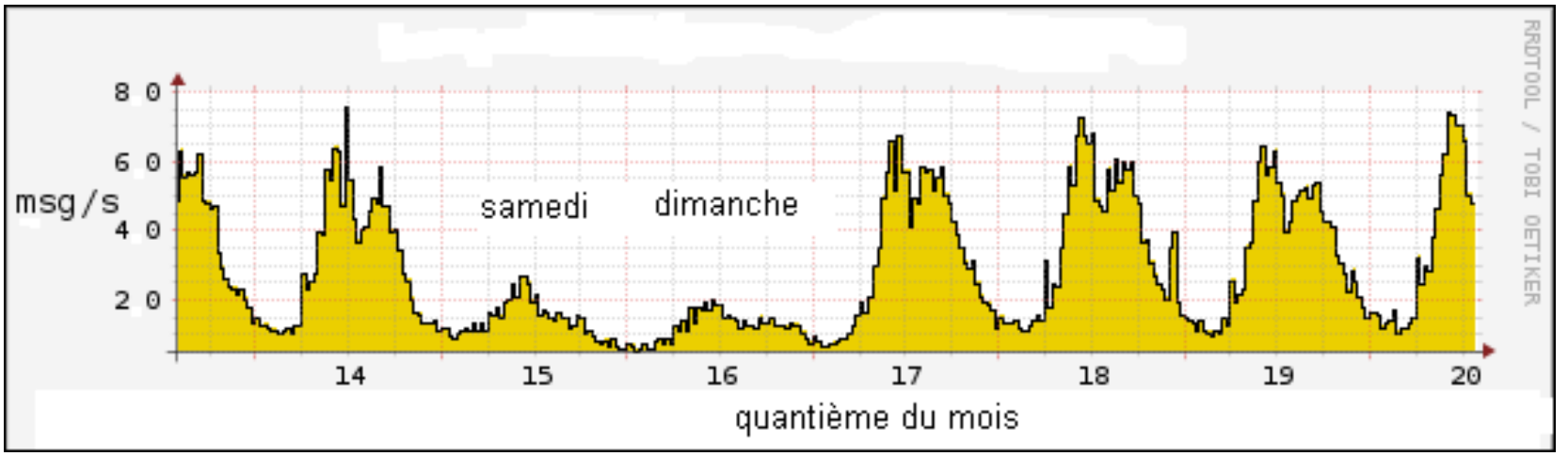
Laurent Aublet-Cuvelier
GIP RENATER
José-Marcio Martins da Cruz
École des Mines de Paris

Sommaire

- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

Service antispam RENATER

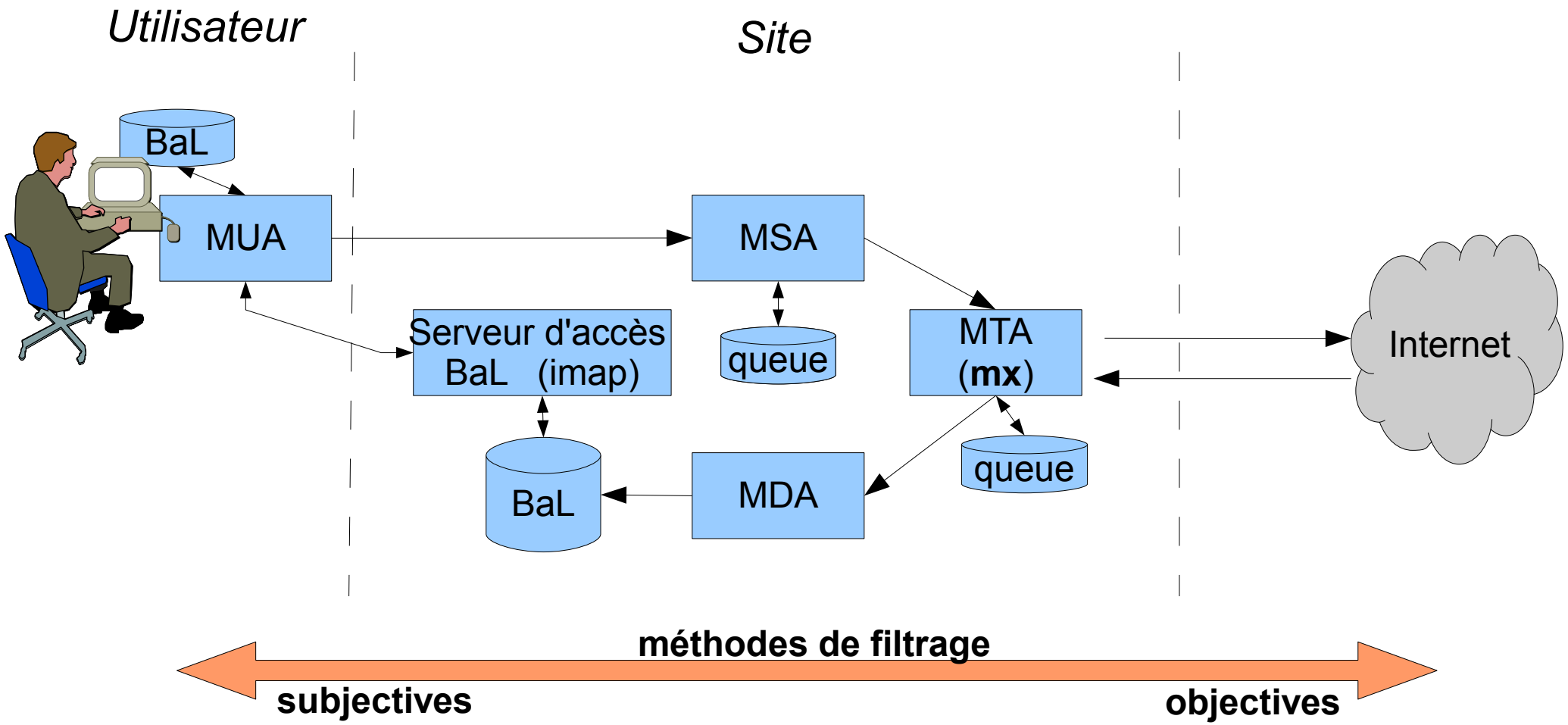
- octobre 2011
  - 47 sites utilisateurs
  - 247 domaines raccordés
  - ~ 600 000 boîtes aux lettres
  - ~ 2 000 000 de messages traités par jour



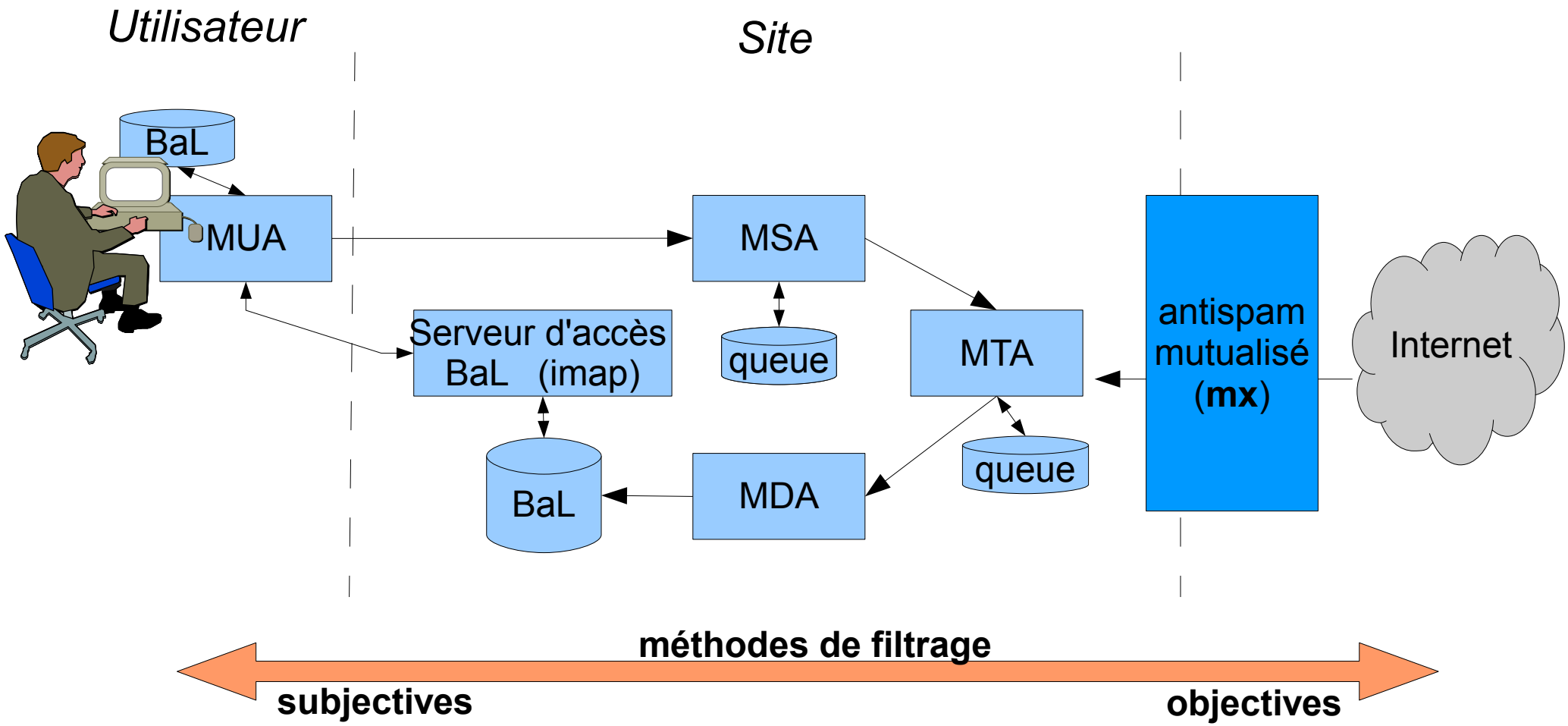
Sommaire

- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

Service antispam et mutualistion

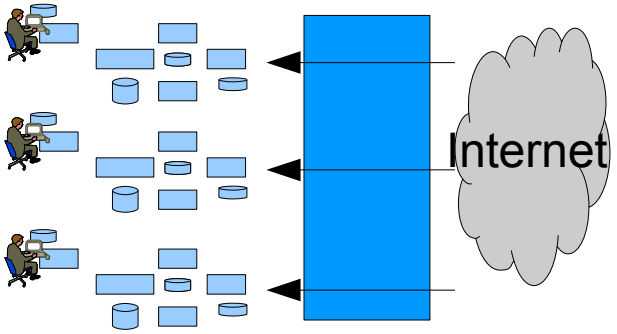


Service antispam et mutualistion



Service antispam et mutualistion

- Spécifités :
  - Distance entre sytème de filtrage et l'utilisateur
  - Contexte enseignement/recherche
    - pas tout à fait un contexte grand public
    - pas tout à fait un contexte entreprise
  - Infrastructure de gestion des boîtes aux lettres
    - par les sites



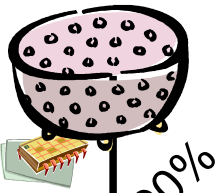
    - plates-formes du marché => plutôt centralisées

Sommaire

- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

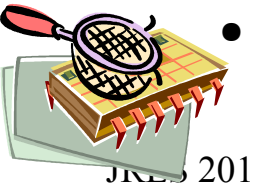
Service antispam : méthodes

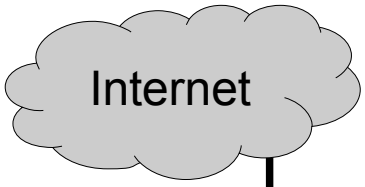
- Typologie des méthodes
  - listes noires et blanches

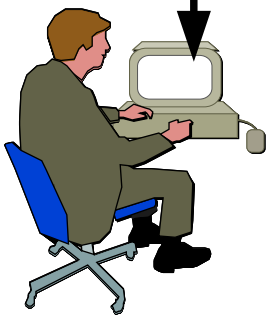


70-90%

- Sur la source
  - listes de réputation
  - Vérification (voire authentification) de la légitimité de la source
- Analyse comportementale
  - Nombre de messages émis, de destinataires, etc.
  - Respect des protocoles, des normes, etc.
- Analyse de contenu



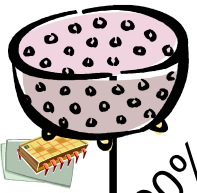




Service antispam : méthodes

- Typologie des méthodes
  - listes noires et blanches

listes B&N statiques

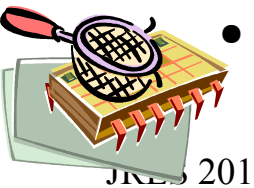


70-90%

  - Sur la source
    - listes de réputation

listes noires dynamiques

    - Vérification (voire authentification) de la légitimité de la source
  - Analyse comportementale
    - Nombre de messages émis, de destinataires, etc.
    - Respect des protocoles, des normes, etc.



  - Analyse de contenu

interaction utilisateur

Sommaire

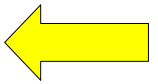
- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

Listes blanches et noires statiques

- Globale et/ou spécifique par domaine
- Rôle correctif
  - bloquer/autoriser une source
    - avant apprentissage par les mécanismes "normaux"
      - par ex. bloquer une source de phishing (avant la RBL)
      - par ex. autoriser un expéditeur anormalement filtré
- Rôle préventif
  - autoriser une source de confiance
    - diffusion CERT, etc.

Listes blanches et noires statiques

- Améliorations
  - accès direct à la modification par les administrateurs de domaines
    - interface web
  - Mutualisation
    - détection des cas d'intérêt général
      - cf. vote électronique
  - Détection de listes blanches contreproductives
    - une liste blanche trop permissive contourne les filtres !
    - Ex. un serveur "blanchi" alors qu'il n'est pas correctement protégé (ex. liste-request...)



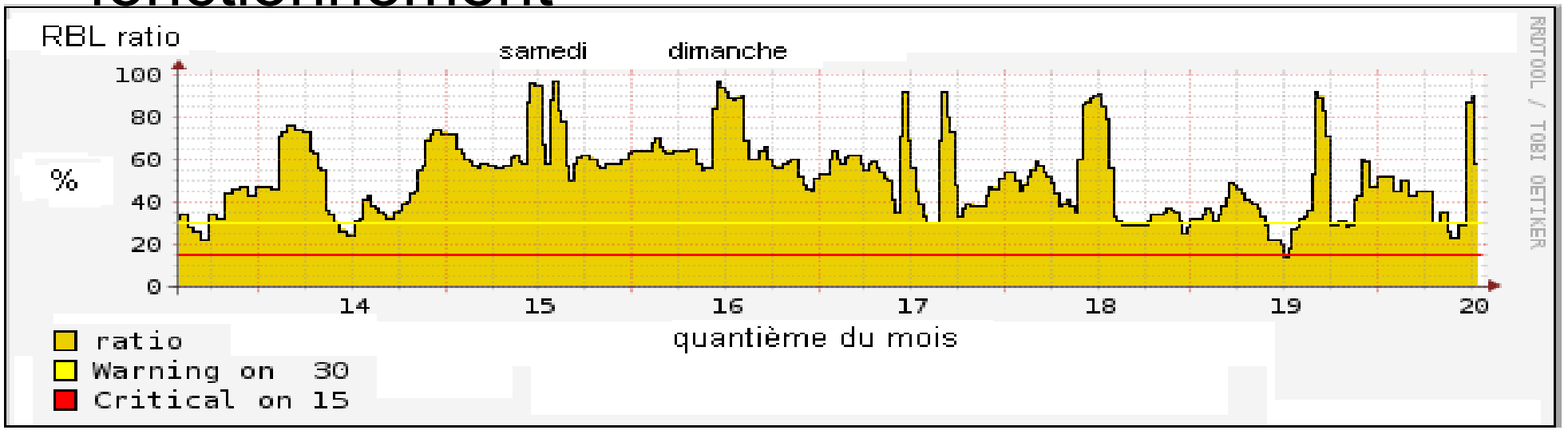
automatisation ?

Sommaire

- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

Listes noires dynamiques (RBL)

- Un choix important : confiance en la source
- Assure un taux de filtrage important
  - 70-90% des rejets
  - => invisibles à l'utilisateur (et l'administrateur)
- Supervision : difficile de détecter un défaut de fonctionnement



Sommaire

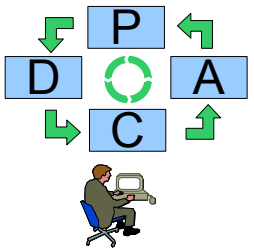
- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

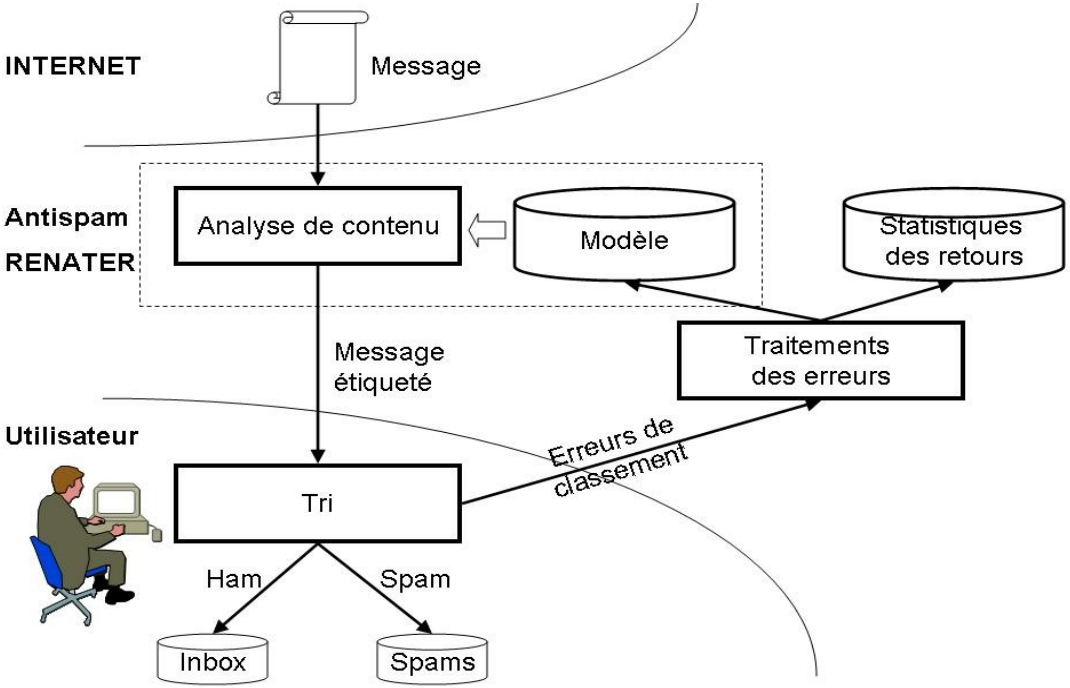
Interaction avec l'utilisateur

- Subjectivité de l'utilisateur
  - Qualité ressentie
    - 70-90% des rejets invisibles...
    - Qualité des 10-30% très visible !
    - *Sensibilité aux faux positifs*
      - *un seul courriel vous manque et tout est dépeuplé...*
  - Notion même de spam diffère...

Interaction avec l'utilisateur

- C'est nécessaire !
  - Besoin d'alimenter le moteur d'analyse de contenu
    - quelle que soit la technique (heuristiques, statistiques, etc.)
  - Besoin de mesure *plus objective* de la qualité
- Boucle de rétroaction





Interaction avec l'utilisateur

- Améliorations :
  - remontées des erreurs de classement
    - via l'adminstrateur
    - via le plug-in Thunderbird (cf. Poster JRES 2011)
  - traitement automatisé
    - analyse des signalements :
      - réelle erreur
      - liste blanche contre-productive
      - hors antispam RENATER

Sommaire

- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

Perspectives

- Test d'un autre moteur d'analyse de contenu
  - "branché" en parallèle du moteur en exploitation
  - test en cours
  - l'analyse s'appuie en partie sur la boucle de rétroaction :
    - les messages mal classés par le moteur A, étaient-ils mieux classés par B ?
    - mais il reste difficile de mesurer les faux positifs de B !
      - sollicitation des administrateurs

Sommaire

- Introduction
- Service antispam et mutualisation
- Service antispam : méthodes
  - Listes blanches et noires statiques
  - Listes noires dynamiques (RBL)
  - Interaction avec l'utilisateur
- Perspectives
- Conclusion

Conclusion

- 100 fois sur le métier...
  - étude solutions alternatives
  - extension du portail web d'administration
  - automatisation de la mutualisation de listes blanches
  - ...
- Objectifs
  - améliorer la qualité du service
  - augmenter la valeur ajoutée tirée de la mutualisation du service