Sauvegarde coopérative

Modèle théorique et solution

Damien Martin-Guillerez

IRISA - ACES

Introduction

- Sauvegarde coopérative.
  - Terminaux mobiles.
  - Possédant un moyen de communication quelconque.
  - Accédant à Internet par intermittance.

Introduction

- Raisons d'une telle sauvegarde.
  - Coût en énergie, en prix d'accès de certains réseaux : sauvegarde systématique sur Internet pas toujours possible ;
  - Solutions usuelles chères (duplication de disque, communications énormes...) ;
  - Duplication sur un autre PDA $\Rightarrow$ possibilité de le revoir ou qu'il sauvegarde sur Internet.

Propriétés

- Appareils mobiles susceptibles de pannes.
- Moyen de communication ad hoc.
- Mémoire de sauvegarde (RAM ou flash).
- Serveurs de sauvegarde sur Internet.

Limitations

- Pas de propagation de l'information.
  - Coût excessif.
  - Gain relatif faible.
- Taille de la mémoire allouée limitée.

Limitations

- Pas de propagation de l'information.
  - Coût excessif.
  - Gain relatif faible.
- Taille de la mémoire allouée limitée.

$\Rightarrow$ Sélection des PDA de sauvegarde.

Modèle

- Fichier de taille k.
- On note les PDA : PDA_i avec PDA_0 représentant Internet.

Modèle

- Fichier de taille k.
- On note les PDA : PDA_i avec PDA_0 représentant Internet.
  - Chaque PDA possède ses propres paramètres : taille disponible, énergies de transfert et de conservation, probabilité de panne, etc...

Modèle

- Fichier de taille k.
- On note les PDA : PDA_i avec PDA_0 représentant Internet.
  - Chaque PDA possède ses propres paramètres : taille disponible, énergies de transfert et de conservation, probabilité de panne, etc...
  - $\mathcal{A}_{i,j}$: évènement de transfert réussi vers Internet ;
  - $\mathcal{B}_{i,j}$: évènement d'aller-retour réussi en passant par le PDA_j.

Modèle

On souhaite pouvoir comparer les PDA accessibles entre eux. Pour cela on comparera la valeur :

$$V_{i,j} = \alpha \, \mathbb{P}_{\mathcal{A}_{i,j}} + (1-\alpha) \, \mathbb{P}_{\mathcal{B}_{i,j}} - \beta \, E_{i,j}$$

Modèle

On souhaite pouvoir comparer les PDA accessibles entre eux. Pour cela on comparera la valeur :

$$V_{i,j} = \alpha \, \mathbb{P}_{\mathcal{A}_{i,j}} + (1 - \alpha) \, \mathbb{P}_{\mathcal{B}_{i,j}} - \beta \, E_{i,j}$$

- α permet de spécifier la priorité donnée à la sauvegarde sur Internet par rapport à la récupération directement sur le PDA ;

Modèle

On souhaite pouvoir comparer les PDA accessibles entre eux. Pour cela on comparera la valeur :

$$V_{i,j} = \alpha\, \mathbb{P}_{\mathcal{A}_{i,j}} + (1-\alpha)\, \mathbb{P}_{\mathcal{B}_{i,j}} - \beta\, E_{i,j}$$

- α permet de spécifier la priorité donnée à la sauvegarde sur Internet par rapport à la récupération directement sur le PDA ;
- β sert à mettre plus ou moins l'accent sur la réduction des coûts énergétique.

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \mathbb{P}_{\text{atteindre Internet}}$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \mathbb{P}_{\text{transmettre le fichier à j}} \cdot \mathbb{P}_{\text{transmission sur Internet}}$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k, 0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \mathbb{P}_{\text{retour du fichier}}$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k, 0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \mathbb{P}_{\text{transmettre le fichier à j}} \cdot \mathbb{P}_{\text{retransmettre le fichier à i}}$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,i}(k, t_{min}, t_{max})$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,i}(k, t_{min}, t_{max})$
- $E_{i,j} =$
  énergie maximale dépensée par la sauvegarde sur j

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,i}(k, t_{min}, t_{max})$
- $E_{i,j} = \max(E_{\mathcal{A}_{i,j}}, E_{\mathcal{B}_{i,j}})$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,i}(k, t_{min}, t_{max})$
- $E_{i,j} = E_{\mathcal{B}_{i,j}}$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,i}(k, t_{min}, t_{max})$
- $E_{i,j} = 2(E^i_{trans}(k) + E^j_{trans}(k)) + E_j(k, t_{max})$

Précisions sur le modèle (1/2)

- $\mathbb{P}_{\mathcal{A}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $\mathbb{P}_{\mathcal{B}_{i,j}} = \tilde{p}_{i,j}(k,0) \cdot \hat{p}_{j,i}(k, t_{min}, t_{max})$
- $E_{i,j} = 2E^{j}_{trans}(k) + E_j(k, t_{max})$

Précisions sur le modèle (2/2)

- $p_{i,j}(t) = \mathbb{P}_{\text{i et j se voient à l'instant } t}$

Précisions sur le modèle (2/2)

- $p_{i,j}(t) = (1 - p^i_{panne}(t)) \cdot (1 - p^j_{panne}(t)) \cdot p'_{i,j}(t)$

Précisions sur le modèle (2/2)

- $p_{i,j}(t) = (1 - p^i_{panne}(t)) \cdot (1 - p^j_{panne}(t)) \cdot p'_{i,j}(t)$
- $\tilde{p}_{i,j}(k,t) = \mathbb{P}_{\text{transmission du fichier à l'instant t}}$

Précisions sur le modèle (2/2)

- $p_{i,j}(t) = (1 - p^i_{panne}(t)) \cdot (1 - p^j_{panne}(t)) \cdot p'_{i,j}(t)$
- $\tilde{p}_{i,j}(k,t) = \min_{t' \in [t, t+t_{trans(k)}]} p_{i,j}(t')$

Précisions sur le modèle (2/2)

- $p_{i,j}(t) = (1 - p^i_{panne}(t)) \cdot (1 - p^j_{panne}(t)) \cdot p'_{i,j}(t)$
- $\tilde{p}_{i,j}(k,t) = \min_{t' \in [t, t + t_{trans(k)}]} p_{i,j}(t')$
- $\hat{p}_{i,j}(k, t_1, t_2) = \mathbb{P}_{\text{transmission du fichier à un instant } t' \in [t_1, t_2]}$

Précisions sur le modèle (2/2)

- $p_{i,j}(t) = (1 - p^i_{panne}(t)) \cdot (1 - p^j_{panne}(t)) \cdot p'_{i,j}(t)$
- $\tilde{p}_{i,j}(k,t) = \min_{t' \in [t, t+t_{trans(k)}]} p_{i,j}(t')$
- $\hat{p}_{i,j}(k,t_1,t_2) = \max_{t' \in [t_1, t_2 - t_{trans(k)}]} \tilde{p}_{i,j}(k,t')$

Paramètres dépendant de j

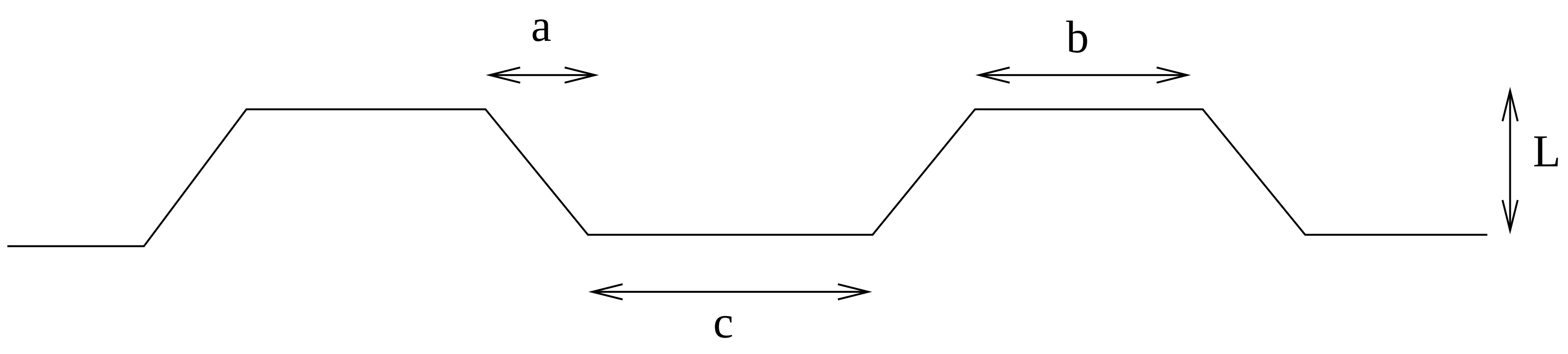
- $a_j(k, t_{max}) = \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $p^j_{panne}(t) = \min(1, b_j\ t)$
- $c_j(k, t_{max}) = 2E^j_{trans}(k) + E_j(k, t_{max})$

Paramètres dépendant de j

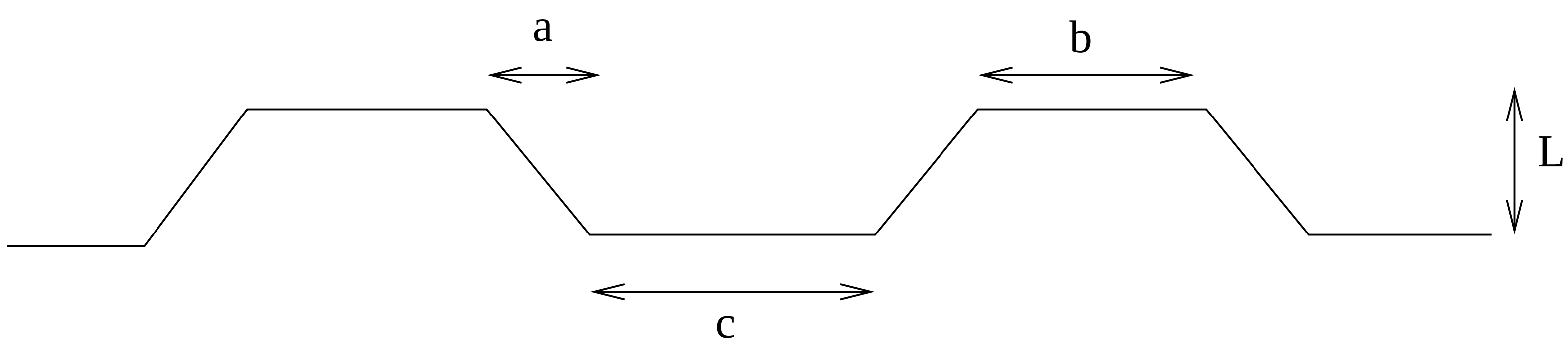
- $a_j(k, t_{max}) = \hat{p}_{j,0}(k, t_{trans}(k), t_{max})$
- $p^j_{panne}(t) = \min(1, b_j\ t)$
- $c_j(k, t_{max}) = 2E^j_{trans}(k) + E_j(k, t_{max})$

$\Rightarrow \quad V_{i,j}(k) = \tilde{p}_{i,j}(k, 0)(\alpha \quad a_j(k) + (1 - \alpha)\hat{p}_{i,j}(k, t_{min}, t_{max})) - \beta\ c_j(k)$

Fonction $p'_{i,j}$

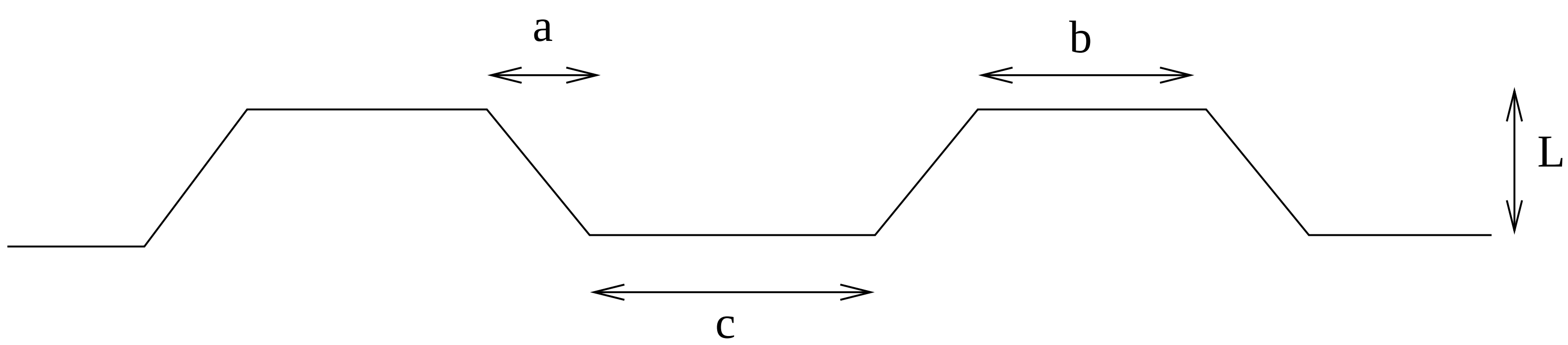


Fonction $p'_{i,j}$



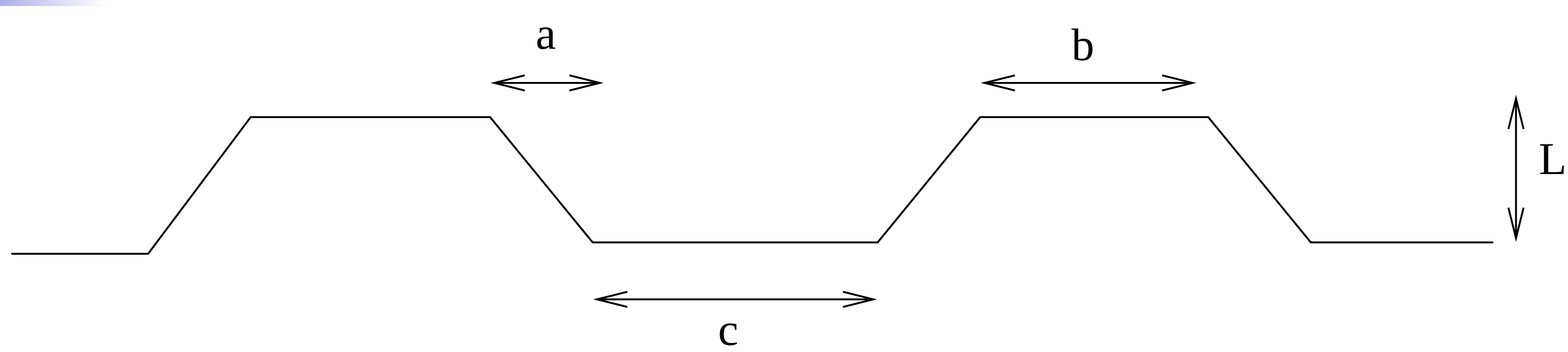
- $n_{seen}^{i,j}$: nombre de fois que i a croisé j ;
- $t_{min}^{i,j}$, $t_{max}^{i,j}$ et $t_{moy}^{i,j}$: temps de visibilité ;
- $t_{lastseen}$: temps depuis la dernière communication entre i et j ;
- $t_{moyseen}$: temps moyen ;

Fonction $p'_{i,j}$



- $L = \min(1, \frac{n_{seen}^{i,j}}{n_{par}})$
- $\epsilon = \frac{t_{min}^{i,j} + t_{max}^{i,j}}{2} - t_{moy}^{i,j}$
- $a = t_{moy}^{i,j} - t_{min}^{i,j} + \epsilon$
- $b = t_{moy}^{i,j} - a$
- $c = t_{moyseen} - a$

Fonction $p'_{i,j}$



Algorithme :

- À la connexion :
  - $t_{moyseen} \leftarrow \frac{t_{lastseen} + n_{seen}^{i,j} t_{moyseen}}{n_{seen}^{i,j} + 1}$.
- À la déconnexion :
  - $t_{max}^{i,j} \leftarrow \max(t_{max}^{i,j}, t_{new}^{i,j})$, $t_{min}^{i,j} \leftarrow \min(t_{min}^{i,j}, t_{new}^{i,j})$ et $t_{moy}^{i,j} \leftarrow \frac{n_{seen}^{i,j} t_{moy}^{i,j} + t_{new}^{i,j}}{n_{seen}^{i,j} + 1}$;
  - $t_{lastseen} \leftarrow 0$ et $n_{seen}^{i,j} \leftarrow n_{seen}^{i,j} + 1$.

Algorithmes

- PDA_i : **Découverte**(k, t_{max})

Algorithmes

- PDA_i : **Découverte**(k, t_{max})
- PDA_j : **Réponse_Découverte**($a_j(k, t_{max}), b_j, c_j(k, t_{max})$) si $k < T_{dispo}^j$

Algorithmes

- PDA_i : **Découverte**(k, t_{max})
- PDA_j : **Réponse_Découverte**($a_j(k, t_{max}), b_j, c_j(k, t_{max})$) si $k < T^j_{dispo}$
- PDA_i : $\forall j$
  - $j = 0 \Rightarrow$ Sauvegarde sur Internet.
  - $j \neq 0 \Rightarrow$ Sauvegarde sur les N PDA_j tel que $V_{i,j}(k)$ soit maximisé.

Quelques précisions...

- On prendra $t_{min} = c - t_{lastseen}$ et $t_{max} = 2c + 4a + 2b$ par rapport à Internet ;
- Si $b > t_{trans(k)}$ alors $\tilde{p}_{i,j}(k,t) = \min_{t' \in [t, t+t_{trans(k)}]} p_{i,j}(t') = \min(p_{i,j}(t), p_{i,j}(t + t_{trans(k)}))$;
- Possibilité de suppression du fichier après t_{max}.

Conclusion

- Une solution qu'il faut tester...
- ...Et dont les différentes valeurs (énergies, temps, etc...) doivent être déterminées ;
- La fonction oracle $p'_{i,j}$ est-elle pertinente ?
- Peut on adapter son calcul au cas réel (intermittent) ?

Questions ?