Detección de nodos maliciosos en procesos de consenso

Miguel Rebollo[1,2], Juan Carlos Losada[1], Rosa Mª Benito[1] and Javier Galeano[1]
[1] Grupo de Sistemas Complejos.
Universidad Politécnica de Madrid, c/ Ramiro de Maeztu, 28040 Madrid
[2] Grupo de Tecnología Informática - Inteligencia Artificial.
Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia

Dada un red de nodos interconectados entre sí, cada uno de los cuales tiene un valor inicial dado, los procesos de consenso permiten, mediante el intercambio de información con los vecinos inmediatos, calcular el valor global de ciertas funciones agregadas [1]. Se trata de un proceso de difusión de información regulado por la siguiente ecuación:

$$x_i(t+1) = x_i(t) + \varepsilon \sum_{j \in N_i} [x_j(t) - x_i(t)]$$

de modo que el valor de cada nodo de la red en una iteración t es una combinación lineal de su propio valor y de los valores de sus vecinos en la iteración anterior. El valor

Sin embargo, para que el proceso funcione, es preciso que todos los nodos se comporten de la misma manera. SI alguno de los nodos introduce alguna alteración, el valor final se verá modificado. Por ejemplo, si un nodo mantiene un valor $x_i(t) = cte.\ \forall t$ el resultado final es que la red entera convergerá a ese valor de x_i. Estas alteraciones pueden ser deliberadas, con el propósito de producir desviaciones interesadas, o accidentales, debido a fallos en alguno de los nodos o a ruido que afecte a los valores intercambiadas durante el proceso entre otros motivos.

En el presente trabajo se estudia un mecanismo que permitiría a una red detectar estas situaciones e incluso corregirlas, de manera que aquellos nodos que realizan correctamente el proceso de consenso alcancen el valor real. La única condición necesaria es que en la primera iteracción se intercambien los valores iniciales reales. A partir de ese momento, el algoritmo es capaz de detectar la desviación en el valor de uno de los vecinos. Pero además de detectar el malfuncionamiento de un nodo, es neceario poder corregir la situación y recuperar el valor de consenso original. Por eso, se proponen dos soluciones. La primera consiste en continuar con el proceso y descontar las variaciones al final, obteniendo el valor de consenso real una vez que se alcanza la convergencia de la solución. Una segunda modificación permitiría al nodo que detecta la desviación descontarla en la propia iteración, evitando que el error se propague al resto de la red.

La formulación del modelo de consenso con nodos maliciosos queda regida por la siguiente ecuación

$$x_i(t+1) = x_i(t) + \varepsilon \sum_{j \in N_i} [x_j(t) - x_i(t)] + u_i(t)$$

donde $u_i(t)$ denota un error aditivo en el instante t. Un nodo k se considera malicioso (o defectuoso) si $u_k(t) > 0$ al menos en un instante t.

Agradecimientos Este trabajo ha sido financiado por el Ministerio de Economía, Industria y Competitividad bajo el proyecto con Nº de expediente TIN2015-65515-C4-1-R.

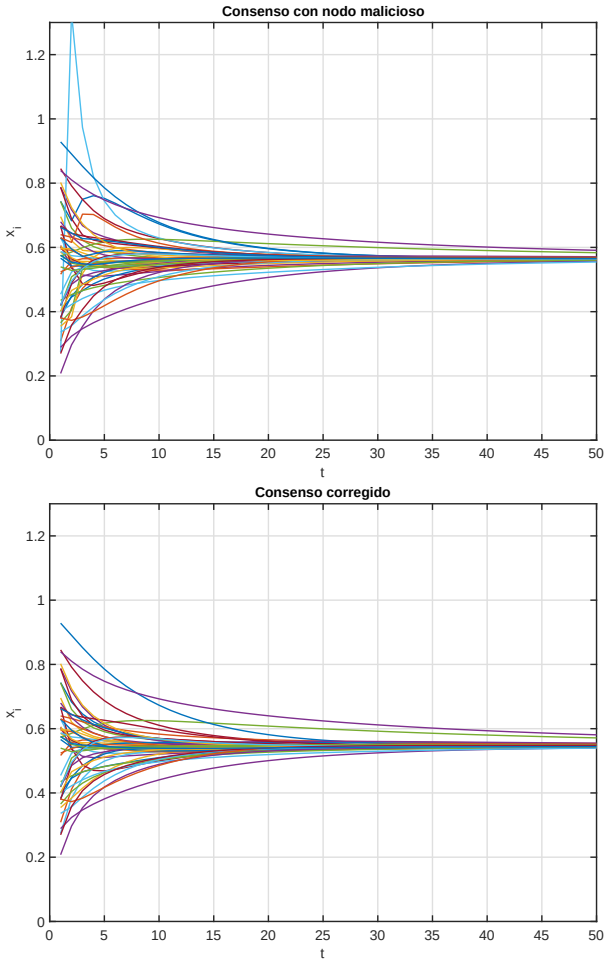

Figure 1: Consenso con un nodo malicioso y consenso corregido durante el proceso. La media real de los valores iniciales es $\bar{x} = 0.5465$ y tras la alteración del nodo pasa a ser $\bar{x'} = 0.5654$, pero la desviación se corrige en la iteración en la que es detectada y la red converge normalmente.

[1] R. Olfati-Saber and R. M. Murray, *Consensus problems in networks of agents with switching topology and time-delays* IEEE TAC. **49**(9), 1520–1533 (2004).