Opinion Formation under Bounded Confidence via Gossip Algorithms and Applications

Nguyen Thi Hoai Linh\textsuperscript{1}, Takayuki Wada\textsuperscript{1}, Izumi Masubuchi\textsuperscript{2}, Toru Asai\textsuperscript{3}, Yasumasa Fujisaki\textsuperscript{1}

\textsuperscript{1}Osaka University, \textsuperscript{2}Kobe University, \textsuperscript{3}Nagoya University
This work is supported by CREST, JST

1 Motivation & Purpose
2 Constant Bounded Confidence Algorithm
3 Clustering Convergence Analysis
4 Numerical simulations
5 Increasing bounded confidence algorithm
6 Applications
Motivation: Public opinion formation
Purpose: Study dynamics of opinion forming in social networks

Considering Problem
We consider a network consisting of $n$ agents. At each time instant agents hold continuous opinions about some issue: $x_i(k)$, $i = 1, 2, \ldots, n$, $k \in \mathbb{N}$.

How agents adjust their opinions while interacting with each other? Which dynamics can result in different formation: Consensus, partial consensus, polarization, fragmentation?
Literature review

Variety of pattern formation → Stubborn Agents

Bounded confidence algorithms

Hegselmann-Krause model

Deffuant-Weisbuch model

\[ x_i(k + 1) = \frac{1}{|\mathcal{N}_i(k)|} x_j(k), \quad i \in V, \]
\[ \mathcal{N}_i(k, d) = \{ j \in V : |x_i(k) - x_j(k)| \leq d \} \]

Chose \((i, j)\) at \(k\),

\[ x_i(k + 1) = \alpha x_i(k) + (1 - \alpha) x_j(k) \]
\[ x_j(k + 1) = \alpha x_j(k) + (1 - \alpha) x_i(k) \]
if \(|x_i(k) - x_j(k)| \leq d\).
2. Constant Bounded Confidence Algorithm

\[ n \text{-agent network, } V = \{1, 2, \ldots, n\} \]

\[ x_i(0) \in \mathbb{R}, \ i \in V: \text{ initial opinion, } \]

\[ d: \text{ confidence threshold} \]

\[ k = 0 \]

\[ k = k + 1 \]

Choose \((i, j)\) with \( p = \frac{2}{n(n-1)} \)

\[ |x_i(k) - x_j(k)| \leq d? \]

Yes:

\[ x_i(k+1) = x_j(k+1) = \frac{x_i(k) + x_j(k)}{2}, \quad x_l(k+1) = x_l(k), \ l \in V \setminus \{i, j\} \]

No:

\[ x_i(k+1) = x_i(k), \quad l \in V \]
2. Constant Bounded Confidence Algorithm

Algorithm in compact form

\[ x(k + 1) = W(i(k), j(k), x(k))x(k). \] (1)

Here

\[ x(k) = [x_1(k), x_2(k), \cdots, x_n(k)]^T \]

\[ W(i, j, x(k)) = \begin{cases} W_{ij}, & \text{if } |x_i(k) - x_j(k)| \leq d, \\ I, & \text{otherwise.} \end{cases} \]

\[ W_{ij} = I - \frac{1}{2}(e_i - e_j)(e_i - e_j)^T. \]
3. Clustering Convergence

Lemma 1 (Equilibrium point)
A vector $x^*$ is an equilibrium point of (1), i.e.,

$$x^* = W(i, j, x^*)x^*, \quad \forall i, j \in V$$

(2)

if and only if it has the form

$$x^*_i = x^*_j \text{ or } |x^*_i - x^*_j| > d \quad \forall i, j \in V.$$  

(3)

Lemma 3

Given $\varepsilon_0 = d/(n - 1)$,

$$\mathbb{P}(\Omega) = 1,$$

$$\Omega = \bigcup_{k_0 \in \mathbb{N}} \left\{ \exists \bar{k} \leq k_0 \text{ s. t. } \forall i, j \in V, \right. \left. \left|x_i(\bar{k}) - x_j(\bar{k})\right| \leq \varepsilon_0 \text{ or } \left|x_i(\bar{k}) - x_j(\bar{k})\right| > d \right\}. $$
3. Clustering Convergence

For each \( x(k) \), we equip the network with a graph

\[ G(x(k)) = (V, E(k)), \]

where

\[ (i, j) \in E(k) \iff |x_i(k) - x_j(k)| \leq \varepsilon_0. \]

**Definition 1. (\( d \)-clusters)**

The clusters \( V_1(k), V_2(k), \ldots, V_{G_d}(k) \) induced from connected components of \( G(x(k)) \) are \( d \)-clusters if

(i) \( \text{dia } V_p(k) := \max_{i, j \in V_p(k)} |x_i(k) - x_j(k)| \leq d \quad \forall p \), and

(ii) \( \text{dist} (V_p(k) - V_q(k)) := \min_{i \in V_p(k), j \in V_q(k)} |x_i(k) - x_j(k)| > d \quad \forall p \neq q. \)
3. Clustering Convergence

Theorem 4.
For any initial opinion profile and a given confidence threshold, with probability one, the opinion profile is partitioned into $d$-clusters within finite step.

Theorem 5.
Consider $d$-cluster $V_1(0)$ with $|V_1(0)| \leq d$. The constant bounded confidence algorithm drives the opinions of all agents to their average $(1^T x(0))/m$ almost surely.
3. Clustering Convergence

**Definition 2.** (Clustering convergence)

Given a parameter $d > 0$. An algorithm given by

$$x(k + 1) = f(x(k)) \quad \text{for } k = 0, 1, 2, \ldots$$

is said to achieve $d$-clustering convergence if starting from any initial state, the system eventually converges to some state $x^*$ of form (3), i.e.,

$$\lim_{k \to \infty} x(k) = x^*.$$

**Theorem 6**

Given a confidence threshold $d$. For any initial opinion profile, Algorithm1 achieves $d$-clustering convergence almost surely.

**Theorem 8.** (Upper bound of clusters)

Given $x(0) \in [0, 1]^n$, $c \in \{1, 2, \ldots, n\}$, if $d \geq 1/c$, then for any sample path of choosing interaction pairs, the number of clusters formed is at most $c$. 

Nguyen Thi Hoai Linh et al.
COOP Workshop, Osaka
4. Numerical Simulation

\[ n = 100, \ x = i(0) \in (0, 1) \text{ uniformly.} \]

\[ d = 0.14 \quad \text{and} \quad d = 1 \]
5. Increasing Bounded Confidence Algorithm

\( n \)-agent network, \( V = \{1, 2, \ldots, n\} \)

\[ c^* : \text{desired number of clusters}, \]
\[ x^* = x(0), \]
\[ d = 0, \delta d : \text{threshold's increment}, \]
\[ \kappa : \text{maximum number of steps} \]

Supervisor calculate \( c(x^*) \):
number of clusters in \( x^* \)

\[ c(x^*) > c^*? \]

- no \( \rightarrow \) Stop
- yes \( \rightarrow \)

\[ x^* = x(\kappa) \]

Constant bounded confident algorithm with
\[ x(0) = x^*, \]
\[ d = d + \delta d, \kappa \]
Theorem 10.

$c(l)$, which is the number of clusters in $x^*(l)$, is a non-increasing function of $l$ with high probability.

Example

$\delta d = 0.001$, $c^* = 5$. Final confidence threshold $d = 0.14$. 

Figure: Increasing confidence threshold with increment 0.001.
We consider the applications of the algorithms in the problem of clustering multi-dimensional data and intrusion detection.

**Clustering Problems**
Given a set of data points, usually in high-dimensional space, partition them into clusters so that:

- Points within each cluster are similar to each other.
- Points from different clusters are dissimilar.
6. Applications

$n$ data objects each of which has $d$ attributes
↔ Multi-agent system $\mathcal{V} = \{1, 2, \ldots, n\}$, $\mathbf{x}_i(0) \in \mathbb{R}^d$

\[
X(k) := \begin{bmatrix}
\mathbf{x}_1(k) & \mathbf{x}_2(k) & \cdots & \mathbf{x}_n(k)
\end{bmatrix}^T \in \mathbb{R}^{n,d}
\]

\[
W(i, j, X(k)) := \begin{cases}
I - \frac{1}{2}(e_i - e_j)(e_i - e_j)^T, & \text{if } \|X_i(k) - X_j(k)\| \leq \delta, \\
I & \text{otherwise.}
\end{cases}
\]
Constant confidence threshold ⇒ Core for clustering-based distributed outlier/anomaly detection in power distribution systems.

Increasing confidence threshold
7. Conclusions

We have studied two algorithms for the dynamics of opinion formation and given a rigorous convergence analysis. The algorithms are

- distributed
- asynchronous
- gossip

The algorithms express the variety of opinion formation: consensus, clustering, fragmentation which are observed frequently in social network.

- have potential to be used as cores of clustering-based distributed outlier/anomaly detection protocol in networks or computer systems, such as power distribution systems.