DenseAlert: Incremental Dense-Subtensor Detection in Tensor Streams - Supplementary Document

KIJUNG SHIN, BRYAN HOOI, JISU KIM and CHRISTOS FALOUTSOS, Carnegie Mellon University

In this supplementary document, we provide properties of the density measure used in the main paper, proofs, and additional experimental results, all of which supplement the main paper [Shin et al. 2017].

ACM Reference Format:
Kijung Shin, Bryan Hooi, and Christos Faloutsos, 2017. DenseAlert: Incremental Dense-Subtensor Detection in Tensor Streams - Supplementary Document *ACM* V, N, Article A (January YYYY), 7 pages.
DOI: http://dx.doi.org/10.1145/0000000.0000000

A. DENSITY MEASURE

In this section, we show that the density measure (Definition 1) used in the main paper satisfies properties that a reasonable "anomalousness" measure should meet. These properties were proposed in [Jiang et al. 2015]. Here, we consider two N-way tensors, $\mathfrak{T}$ of size $I_1 \times I_2 \times ... \times I_N$ and $\mathfrak{T}'$ of size $I'_1 \times I'_2 \times ... \times I'_N$. We denote the sum of the entries in each tensor by $sum(\mathfrak{T})$ and $sum(\mathfrak{T}')$, and define their average entry value as $\bar{t} = \frac{sum(\mathfrak{T})}{\prod_{n=1}^{N} I_N}$ and $\bar{t'} = \frac{sum(\mathfrak{T}')}{\prod_{n=1}^{N} I'_N}$. We first list three basic axioms that any anomalousness measure f should meet.

Axiom 1 (Entry Sum). *Between two tensors with the same dimensionality, one with higher entry sum is more anomalous than the other. Formally, suppose $1 \leq \forall n \leq N$, $I_n = I'_n$. Then $sum(\mathfrak{T}) \geq sum(\mathfrak{T}')$ implies $f(\mathfrak{T}) \geq f(\mathfrak{T}')$, and equality holds if and only if $sum(\mathfrak{T}) = sum(\mathfrak{T}')$.*

Axiom 2 (Concentration). *Between two tensors with the same entry sum, one with smaller dimensionality is more anomalous than the other. Formally, suppose $sum(\mathfrak{T}) = sum(\mathfrak{T}')$. Then $1 \leq \forall n \leq N$, $I_n \leq I'_n$ implies $f(\mathfrak{T}) \geq f(\mathfrak{T}')$, and equality holds if and only if $1 \leq \forall n \leq N$, $I_n = I'_n$.*

Axiom 3 (Size). *Between two tensors with the same average entry value, one with larger dimensionality is more anomalous than the other. Formally, suppose $\bar{t} = \bar{t'}$. Then $1 \leq \forall n \leq N$, $I_n \geq I'_n$ implies $f(\mathfrak{T}) \geq f(\mathfrak{T}')$, and equality holds if and only if $1 \leq \forall n \leq N$, $I_n = I'_n$.*

Although two more axioms were proposed in [Jiang et al. 2015], they are not relevant to our work. The additional axioms are for comparison between subtensors in different tensors, while in our work, only subtensors within the same tensor are compared.

As in the main paper, the densities of $\mathfrak{T}$ and $\mathfrak{T}'$, denoted by $\rho(\mathfrak{T})$ and $\rho(\mathfrak{T}')$, are defined as $\rho(\mathfrak{T}) = \frac{sum(\mathfrak{T})}{\sum_{n=1}^{N} I_N}$ and $\rho(\mathfrak{T}') = \frac{sum(\mathfrak{T}')}{\sum_{n=1}^{N} I'_N}$, respectively. We prove that Axioms 1-3 are satisfied by the density measure used in the paper, in Theorem A.1.

DOI: http://dx.doi.org/10.1145/0000000.0000000

Theorem A.1. *Let $\mathfrak{T}$ and $\mathfrak{T}'$ be two N-way tensors, and let ρ be the density measure in Definition 1. Then ρ satisfies* Axiom 1-3*, i.e.*

(i) *(Entry Sum). Suppose $1 \leq \forall n \leq N$, $I_n = I'_n$. Then $sum(\mathfrak{T}) \geq sum(\mathfrak{T}')$ implies $\rho(\mathfrak{T}) \geq \rho(\mathfrak{T}')$, and equality holds if and only if $sum(\mathfrak{T}) = sum(\mathfrak{T}')$.*

(ii) *(Concentration). Suppose $sum(\mathfrak{T}) = sum(\mathfrak{T}')$. Then $1 \leq \forall n \leq N$, $I_n \leq I'_n$ implies $\rho(\mathfrak{T}) \geq \rho(\mathfrak{T}')$, and equality holds if and only if $1 \leq \forall n \leq N$, $I_n = I'_n$.*

(iii) *(Size). Suppose $\bar{t} = \bar{t}'$. Then $1 \leq \forall n \leq N$, $I_n \geq I'_n$ implies $\rho(\mathfrak{T}) \geq \rho(\mathfrak{T}')$, and equality holds if and only if $1 \leq \forall n \leq N$, $I_n = I'_n$.*

Proof. (i) Suppose $1 \leq \forall n \leq N$, $I_n = I'_n$. Then $sum(\mathfrak{T}) \geq sum(\mathfrak{T}')$ implies

$$\rho(\mathfrak{T}) = \frac{sum(\mathfrak{T})}{\sum_{n=1}^{N} I_n} \geq \frac{sum(\mathfrak{T}')}{\sum_{n=1}^{N} I_n} = \frac{sum(\mathfrak{T}')}{\sum_{n=1}^{N} I'_n} = \rho(\mathfrak{T}').$$

And equivalents for the equality condition are

$$\rho(\mathfrak{T}) = \rho(\mathfrak{T}') \iff \frac{sum(\mathfrak{T})}{\sum_{n=1}^{N} I_n} = \frac{sum(\mathfrak{T}')}{\sum_{n=1}^{N} I'_n} \iff sum(\mathfrak{T}) = sum(\mathfrak{T}').$$

(ii) Suppose $sum(\mathfrak{T}) = sum(\mathfrak{T}')$. Then $1 \leq \forall n \leq N$, $I_n \leq I'_n$ implies

$$\rho(\mathfrak{T}) = \frac{sum(\mathfrak{T})}{\sum_{n=1}^{N} I_n} \geq \frac{sum(\mathfrak{T})}{\sum_{n=1}^{N} I'_n} = \frac{sum(\mathfrak{T}')}{\sum_{n=1}^{N} I'_n} = \rho(\mathfrak{T}').$$

And equivalents for the equality condition are

$$\rho(\mathfrak{T}) = \rho(\mathfrak{T}') \iff \frac{sum(\mathfrak{T})}{\sum_{n=1}^{N} I_n} = \frac{sum(\mathfrak{T}')}{\sum_{n=1}^{N} I'_n} \iff \sum_{n=1}^{N} I_n = \sum_{n=1}^{N} I'_n.$$

Then under $1 \leq \forall n \leq N$, $I_n \leq I'_n$ conditions, $\sum_{n=1}^{N} I_n = \sum_{n=1}^{N} I'_n$ can happen if and only if $1 \leq \forall n \leq N$, $I_n = I'_n$.

(iii) We first prove that $1 \leq \forall n \leq N$, $I_n \geq I'_n$ implies $\frac{\prod_{n=1}^{N} I_n}{\sum_{n=1}^{N} I_n} \geq \frac{\prod_{n=1}^{N} I'_n}{\sum_{n=1}^{N} I'_n}$, and the equality holds if and only if $1 \leq \forall n \leq N$, $I_n = I'_n$. This is since

$$\frac{\prod_{n=1}^{N} I_n}{\sum_{n=1}^{N} I_n} = \frac{1}{\sum_{n=1}^{N} \frac{1}{\prod_{j \neq n} I_j}} \geq \frac{1}{\sum_{n=1}^{N} \frac{1}{\prod_{j \neq n} I'_j}} = \frac{\prod_{n=1}^{N} I'_n}{\sum_{n=1}^{N} I'_n}.$$

And $\frac{\prod_{n=1}^{N} I_n}{\sum_{n=1}^{N} I_n} = \frac{\prod_{n=1}^{N} I'_n}{\sum_{n=1}^{N} I'_n}$ if and only if $\sum_{n=1}^{N} \frac{1}{\prod_{j \neq n} I_j} = \sum_{n=1}^{N} \frac{1}{\prod_{j \neq n} I'_j}$. Under $1 \leq \forall n \leq N, I_n \geq I'_n$ condition, this can happen if and only if $1 \leq \forall n \leq N$, $I_n = I'_n$.

Now, suppose $\bar{t} = \bar{t}'$, i.e. $\frac{sum(\mathfrak{T})}{\prod_{n=1}^{N} I_n} = \frac{sum(\mathfrak{T})}{\prod_{n=1}^{N} I'_n}$. Then, $1 \leq \forall n \leq N$, $I_n \geq I'_n$ implies $\frac{\prod_{n=1}^{N} I_n}{\sum_{n=1}^{N} I_n} \geq \frac{\prod_{n=1}^{N} I'_n}{\sum_{n=1}^{N} I'_n}$, hence

$$\rho(\mathfrak{T}) = \frac{sum(\mathfrak{T})}{\sum_{n=1}^{N} I_n} = \frac{sum(\mathfrak{T})}{\prod_{n=1}^{N} I_n} \times \frac{\prod_{n=1}^{N} I_n}{\sum_{n=1}^{N} I_n} \geq \frac{sum(\mathfrak{T}')}{\prod_{n=1}^{N} I'_n} \times \frac{\prod_{n=1}^{N} I'_n}{\sum_{n=1}^{N} I'_n} = \frac{sum(\mathfrak{T}')}{\sum_{n=1}^{N} I'_n} = \rho(\mathfrak{T}').$$

And equivalents for the equality condition are

$$\rho(\mathfrak{T}) = \rho(\mathfrak{T}') \iff \frac{\prod_{n=1}^{N} I_n}{\sum_{n=1}^{N} I_n} = \frac{\prod_{n=1}^{N} I'_n}{\sum_{n=1}^{N} I'_n} \iff 1 \leq \forall n \leq N,\, I_n = I'_n.$$

■

B. PROOFS

B.1. Lemma B.1

Lemma B.1. *Let $\mathfrak{T}$ be an N-way tensor, and let $\mathfrak{T}'$ be the updated $\mathfrak{T}$ after the change of either $((i_1,...,i_N),\delta,+)$ or $((i_1,...,i_N),\delta,-)$. Let π be a D-ordering of Q, and let $q_f := \arg\min_{q\in C} \pi^{-1}(q)$ where $C = \{(n,i_n) : 1 \leq \forall n \leq N\}$. For $c \in \mathbb{R}_{>0}$, let $M(c)$ be the set of slice indices that are located after q_f in π and having $d_\pi(\cdot)$ at least c, i.e.*

$$M(c) := \{q \in Q : \pi^{-1}(q) > \pi^{-1}(q_f) \wedge d_\pi(q) \geq c\}. \tag{9}$$

And let $j_L, j_H \in [|Q|]$ be satisfying

$$j_H = \begin{cases} \min_{q\in M(c)} \pi^{-1}(q) - 1 & \text{if } M(c) \neq \emptyset, \\ |Q| & \text{otherwise,} \end{cases} \tag{10}$$

and $j_L \leq j_H$. Let $R = \{q \in Q : \pi^{-1}(q) \in [j_L, j_H]\}$. Let π' be an ordering of slice indices Q in $\mathfrak{T}'$ where $\forall j \notin [j_L, j_H]$, $\pi'(j) = \pi(j)$ and $\forall j \in [j_L, j_H]$,

$$d(\mathfrak{T}'(Q_{\pi',\pi'(j)}), \pi'(j)) = \min_{r\in R\cap Q_{\pi',\pi'(j)}} d(\mathfrak{T}'(Q_{\pi',\pi'(j)}), r). \tag{11}$$

Then,

(i) For all $q \in Q$ with $\pi^{-1}(q) \notin [j_L, j_H]$,

$$Q_{\pi,q} = Q_{\pi',q}. \tag{12}$$

(ii) For all $q \in Q$ with $\pi^{-1}(q) > j_H$,

$$\mathfrak{T}(Q_{\pi,q}) = \mathfrak{T}'(Q_{\pi',q}). \tag{13}$$

(iii) For all $q \in Q$ with $j_L \leq \pi^{-1}(q) \leq j_H$,

$$d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r\in R\cap Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r). \tag{14}$$

(iv) For all $q \in Q$ with $j_L \leq \pi^{-1}(q) \leq j_H$,

$$\min_{r\in Q_{\pi',q}-R} d(\mathfrak{T}'(Q_{\pi',q}), r) \geq c. \tag{15}$$

(v) For all $q \in Q$ with $j_L \leq \pi^{-1}(q) \leq j_H$, let $r := \arg\min_{q'\in Q_{\pi',q}} \pi^{-1}(q')$ be the slice index located earliest in π among $Q_{\pi',q}$. Then the following inequalities hold:

$$d(\mathfrak{T}'(Q_{\pi',q}), q) \leq d(\mathfrak{T}'(Q_{\pi',q}), r), \tag{16}$$

$$d(\mathfrak{T}(Q_{\pi',q}), r) \leq d_\pi(r). \tag{17}$$

(vi) For all $q \in Q$ with $\pi^{-1}(q) > j_H$,

$$d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r\in Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r). \tag{18}$$

Proof. (i) If $\pi^{-1}(q) \notin [j_L, j_H]$, then either $\pi^{-1}(q) < j_L$ or $\pi^{-1}(q) > j_H$. Consider $\pi^{-1}(q) < j_L$ first. Since π and π' coincide on $[1, j_L)$, $\pi'^{-1}(q) = \pi^{-1}(q) < j_L$ holds, so π and π' coincide on $[1, \pi^{-1}(q))$ as well. This implies that

$$Q_{\pi,q} = Q \setminus \{r \in Q : \pi^{-1}(r) < \pi^{-1}(q)\} = Q \setminus \{r \in Q : \pi'^{-1}(r) < \pi'^{-1}(q)\} = Q_{\pi',q}.$$

Now consider $\pi^{-1}(q) > j_H$ case. Since π and π' coincide on $(j_H, |Q|]$, $\pi'^{-1}(q) = \pi^{-1}(q) > j_H$ holds, so π and π' coincide on $[\pi^{-1}(q), |Q|]$ as well. This implies that

$$Q_{\pi,q} = \{r \in Q : \pi^{-1}(r) \geq \pi^{-1}(q)\} = \{r \in Q : \pi'^{-1}(r) \geq \pi'^{-1}(q)\} = Q_{\pi',q}.$$

Hence for either cases, Eq. (12) holds.

(ii) Since $\pi^{-1}(q) > j_H$, Lemma B.1 (i) implies $Q_{\pi,q} = Q_{\pi',q}$. And since $\pi^{-1}(q) > j_H \geq \pi^{-1}(q_f) = \min_{q' \in C} \pi^{-1}(q')$, the changed entry $t_{i_1 \ldots i_N}$ with index $(i_1, ..., i_N)$ is not contained in $\mathfrak{T}(Q_{\pi,q})$. These together imply Eq. (13).

(iii) Note that π and π' coinciding on $[|Q|] \backslash [j_L, j_H]$ implies that $j_L \leq \pi'^{-1}(q) \leq j_H$ as well. Hence this with the condition of π' in Eq. (11) implies Eq. (14).

(iv) If $M(c) = \emptyset$, then $Q_{\pi',q} \backslash R = \emptyset$, so there is nothing to show. When $M(c) \neq \emptyset$ so that $Q_{\pi',q} \backslash R \neq \emptyset$, fix any $r \in Q_{\pi',q} \backslash R$, and let $q_h := \pi(j_H + 1)$. We show Eq. (15) as $d(\mathfrak{T}'(Q_{\pi',q}), r) \geq d(\mathfrak{T}'(Q_{\pi',q_h}), r) \geq d_\pi(q_h) \geq c$. First, π and π' coinciding on $[|Q|] \backslash [j_L, j_H]$ implies that $\pi'^{-1}(q) \leq j_H < j_H + 1 = \pi'^{-1}(q_h)$, so $Q_{\pi',q} \supset Q_{\pi',q_h}$. And $\pi'^{-1}(r) = \pi^{-1}(r) \geq j_H + 1 = \pi^{-1}(q_h) = \pi'^{-1}(q_h)$ implies $r \in Q_{\pi',q_h}$, which implies $d(\mathfrak{T}'(Q_{\pi',q}), r) \geq d(\mathfrak{T}'(Q_{\pi',q_h}), r)$. Also, since $\pi^{-1}(q_h) > j_H$, Lemma B.1 (ii) implies $\mathfrak{T}(Q_{\pi,q_h}) = \mathfrak{T}'(Q_{\pi',q_h})$, and hence $d(\mathfrak{T}'(Q_{\pi',q_h}), r) \geq d(\mathfrak{T}(Q_{\pi,q_h}), r)$. Also, π being a D-ordering implies $d(\mathfrak{T}(Q_{\pi,q_h}), r) \geq d(\mathfrak{T}(Q_{\pi,q_h}), q_h) = d_\pi(q_h)$. Also, $M(c) \neq \emptyset$ and $j_H + 1 = \min_{q \in M(c)} \pi^{-1}(q)$ in Eq. (10) implies $q_h \in M(c)$, and definition of $M(c)$ in Eq. (9) implies that $d_\pi(q_h) \geq c$. These together imply Eq. (15).

(v) Note that π and π' coinciding on $[|Q|] \setminus [j_L, j_H]$ with $\pi^{-1}(q) \in [j_L, j_H]$ implies that $\pi^{-1}(r) \in [j_L, j_H]$, i.e. $r \in R$. Then Eq. (16) is from the condition of π' in Eq. (11) and that $r \in R \cap Q_{\pi',q}$. Also, $r = \arg\min_{q' \in Q_{\pi',q}} \pi^{-1}(q')$ implies $Q_{\pi',q} \subset Q_{\pi,r}$, which implies Eq. (17).

(vi) We show Eq. (18) as $d(\mathfrak{T}'(Q_{\pi',q}), q) = d(\mathfrak{T}(Q_{\pi,q}), q) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}(Q_{\pi,q}), r) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$. First, from Lemma B.1 (ii), $d(\mathfrak{T}'(Q_{\pi',q}), q) = d(\mathfrak{T}(Q_{\pi,q}), q)$ holds. Next, since π is a D-ordering, $d(\mathfrak{T}(Q_{\pi,q}), q) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}(Q_{\pi,q}), r)$ holds. Lastly, $\min_{r \in Q_{\pi,q}} d(\mathfrak{T}(Q_{\pi,q}), r) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$ is again from Lemma B.1 (ii). These together imply Eq. (18). ∎

B.2. Proof of Lemma 3.7

In the following proof, we use $q_f := \arg\min_{r \in C} \pi^{-1}(r) = \pi(j_L)$ where $C = \{(n, i_n) : 1 \leq \forall n \leq N\}$, $M := \{q \in Q : \pi^{-1}(q) > \pi^{-1}(q_f) \ \wedge \ d_\pi(q) \geq d_\pi(q_f) + \delta\}$, and $R := \{q \in Q : \pi^{-1}(q) \in [j_L, j_H]\}$, all of which are defined in the main paper. Note that $\mathfrak{T}$, $\mathfrak{T}'$, π, q_f, j_L, j_H, R, π' satisfies the conditions in Lemma B.1 with $M = M(d_\pi(q_f) + \delta)$, and hence Lemma B.1 is applicable.

Proof. From Definition 3.1 of D-ordering, we need to show that for all $q \in Q$,

$$d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r). \tag{19}$$

We divide into 3 cases depending on the location of q with respect to π: (i) $\pi^{-1}(q) < j_L$, (ii) $j_L \leq \pi^{-1}(q) \leq j_H$, and (iii) $\pi^{-1}(q) > j_H$.

(i) For the case of $\pi^{-1}(q) < j_L$, we show Eq. (19) as $d(\mathfrak{T}'(Q_{\pi',q}), q) = d(\mathfrak{T}(Q_{\pi,q}), q) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}(Q_{\pi,q}), r) \leq \min_{r \in Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$. At first, $\pi^{-1}(q) < j_L$ and Lemma B.1 (i) imply

$$Q_{\pi',q} = Q_{\pi,q}. \tag{20}$$

Also, from $\pi^{-1}(q) < j_L = \min_{q' \in C} \pi^{-1}(q')$ where $C = \{(n, i_n) : 1 \leq \forall n \leq N\}$, the changed entry $t_{i_1 \ldots i_N}$ is not in the slice with index q. From this and Eq. (20), $d(\mathfrak{T}'(Q_{\pi',q}), q) = d(\mathfrak{T}(Q_{\pi,q}), q)$

holds. Next, from π being a D-ordering, we have $d(\mathfrak{T}(Q_{\pi,q}), q) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}(Q_{\pi,q}), r)$ holds. Lastly, from Eq. (20) and that the slice sums of the slices in Q either increase or remain the same under $((i_1, ..., i_N), \delta, +)$, $\min_{r \in Q_{\pi,q}} d(\mathfrak{T}(Q_{\pi,q}), r) \leq \min_{r \in Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$ holds. From these, $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$ holds for $\pi^{-1}(q) < j_L$.

(ii) For the case of $j_L \leq \pi^{-1}(q) \leq j_H$, we show Eq. (19) by $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q} \cap R} d(\mathfrak{T}'(Q_{\pi',q}), r)$ and $\min_{r \in Q_{\pi',q} - R} d(\mathfrak{T}'(Q_{\pi',q}), r) \geq d_\pi(q_f) + \delta \geq d(\mathfrak{T}'(Q_{\pi',q}), q)$. At first, Lemma B.1 (iii) implies $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q} \cap R} d(\mathfrak{T}'(Q_{\pi',q}), r)$. Also, since $M = M(d_\pi(q_f) + \delta)$, Lemma B.1 (iv) implies $\min_{r \in Q_{\pi',q} - R} d(\mathfrak{T}'(Q_{\pi',q}), r) \geq d_\pi(q_f) + \delta$. For $d(\mathfrak{T}'(Q_{\pi',q}), q) \leq d_\pi(q_f) + \delta$, let $r := \arg\min_{q' \in Q_{\pi',q}} \pi^{-1}(q')$ be the slice with the index earliest in π among $Q_{\pi',q}$, and we show $d(\mathfrak{T}'(Q_{\pi',q}), q) \leq d(\mathfrak{T}'(Q_{\pi',q}), r) \leq d_\pi(q_f) + \delta$. $d(\mathfrak{T}'(Q_{\pi',q}), q) \leq d(\mathfrak{T}'(Q_{\pi',q}), r)$ is from Eq. (16) in Lemma B.1 (v). For $d(\mathfrak{T}'(Q_{\pi',q}), r) \leq d_\pi(q_f) + \delta$, we divide into cases where $r = q_f$ or $r \neq q_f$. When $r = q_f$, note that slice sums can increase at most δ under $((i_1, ..., i_N), \delta, +)$, so $d(\mathfrak{T}'(Q_{\pi',q}), q_f) \leq d(\mathfrak{T}(Q_{\pi',q}), q_f) + \delta$ holds. And Eq. (17) in Lemma B.1 (v) implies $d(\mathfrak{T}(Q_{\pi',q}), q_f) + \delta \leq d_\pi(q_f) + \delta$. Hence these implies $d(\mathfrak{T}'(Q_{\pi',q}), r) \leq d_\pi(q_f) + \delta$ for $r = q_f$. When $r \neq q_f = \pi(j_L)$, π and π' coinciding on $[1, j_L)$ implies that $q_f \notin Q_{\pi',q}$, which implies that $t_{i_1 \dots i_N}$ is not in $\mathfrak{T}(Q_{\pi',q})$. Hence $\mathfrak{T}(Q_{\pi',q}) = \mathfrak{T}'(Q_{\pi',q})$, and this implies $d(\mathfrak{T}'(Q_{\pi',q}), r) = d(\mathfrak{T}(Q_{\pi',q}), r)$. Then Eq. (17) in Lemma B.1 (v) implies $d(\mathfrak{T}(Q_{\pi',q}), r) \leq d_\pi(r)$. Then, π and π' coinciding on $[|Q|] \backslash [j_L, j_H]$, $\pi^{-1}(q) \leq j_H$, and $r \neq \pi(j_L)$ imply that $j_L < \pi^{-1}(r) \leq \pi^{-1}(q) \leq j_H$, and Eq. (2) implies that $r \notin M$. Hence $d_\pi(r) < d_\pi(q_f) + \delta$ holds, and these imply $d(\mathfrak{T}'(Q_{\pi',q}), r) \leq d_\pi(q_f) + \delta$ for $r \neq q_f$. Hence $d(\mathfrak{T}'(Q_{\pi',q}), r) \leq d_\pi(q_f) + \delta$ holds for any r. This with $d(\mathfrak{T}'(Q_{\pi',q}), q) \leq d(\mathfrak{T}'(Q_{\pi',q}), r)$ implies that for $j_L \leq \pi^{-1}(q) \leq j_H$ we have

$$d(\mathfrak{T}'(Q_{\pi',q}), q) \leq d_\pi(q_f) + \delta. \tag{21}$$

Then Lemma B.1 (iii), (iv), and Eq. (21) imply that $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$ holds for $j_L \leq \pi^{-1}(q) \leq j_H$.

(iii) For the case of $\pi^{-1}(q) > j_H$, Lemma B.1 (vi) implies Eq. (19). ∎

B.3. Proof of Lemma 3.11

In the following proof, we use $q_f := \arg\min_{r \in C} \pi^{-1}(r)$ where $C = \{(n, i_n) : 1 \leq \forall n \leq N\}$, $M_{\min} := \{q \in Q : d_\pi(q) > c_\pi(q) - \delta\}$, $M_{\max} := \{q \in Q : \pi^{-1}(q) > \pi^{-1}(q_f) \ \wedge \ d_\pi(q) \geq c_\pi(q_f)\}$, and $R := \{q \in Q : \pi^{-1}(q) \in [j_L, j_H]\}$, all of which are defined in the main paper. Note that $\mathfrak{T}$, $\mathfrak{T}'$, π, q_f, j_L, j_H, R, π' satisfies the conditions in Lemma B.1 with $M_{\max} = M(c_\pi(q_f))$, and hence Lemma B.1 is applicable.

Proof. From Definition 3.1 of D-ordering, we need to show that for all $q \in Q$,

$$d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q}} d(\mathfrak{T}'(Q_{\pi',q}), r). \tag{22}$$

We divide into 3 cases depending on the location of q with respect to π: (i) $\pi^{-1}(q) < j_L$, (ii) $j_L \leq \pi^{-1}(q) \leq j_H$, and (iii) $\pi^{-1}(q) > j_H$.

(i) For the case of $\pi^{-1}(q) < j_L$, we show as $d(\mathfrak{T}'(Q_{\pi',q}), r) \geq d_\pi(q) = d(\mathfrak{T}'(Q_{\pi',q}), q)$. At first, $\pi^{-1}(q) < j_L$ and Lemma B.1 (i) imply

$$Q_{\pi',q} = Q_{\pi,q}. \tag{23}$$

Also, $j_L \le \pi^{-1}(q_f) = \min_{q' \in C} \pi^{-1}(q')$ where $C = \{(n, i_n) : 1 \le \forall n \le N\}$, so the changed entry $t_{i_1 \dots i_N}$ is not in the slice with index q. This and Eq. (23) imply that $d_\pi(q) = d(\mathfrak{T}(Q_{\pi,q}), q) = d(\mathfrak{T}'(Q_{\pi',q}), q)$ holds. For showing $d(\mathfrak{T}'(Q_{\pi',q}), r) \ge d_\pi(q)$, we divide into cases when $r \in C$ and $r \in Q_{\pi',q} - C$. For $r \in C$ case, let $x \in Q$ be satisfying $\pi^{-1}(x) \le \pi^{-1}(q_f)$ and $d_\pi(x) = c_\pi(q_f)$, whose existence is from the definition of $c_\pi(\cdot)$. We show by $d(\mathfrak{T}'(Q_{\pi',q}), r) + \delta \ge d(\mathfrak{T}(Q_{\pi,q}), r) \ge d(\mathfrak{T}(Q_{\pi,x}), r) \ge d_\pi(x) = c_\pi(q_f) \ge d_\pi(q) + \delta$. Since slice sums decrease at most δ under $((i_1, ..., i_N), \delta, -)$ and from Eq. (23), $d(\mathfrak{T}'(Q_{\pi',q}), r) + \delta \ge d(\mathfrak{T}(Q_{\pi,q}), r)$ holds. Also, $d_\pi(x) = c_\pi(q_f) > c_\pi(q_f) - \delta$ implies that $x \in M_{\min}$ from definition of $M_{\min}$, hence $\pi^{-1}(x) \ge j_L = \min_{q \in M_{\min}} \pi^{-1}(q)$. Then $\pi^{-1}(q) < j_L \le \pi^{-1}(x)$ implies $Q_{\pi,x} \subset Q_{\pi,q}$ and $\pi^{-1}(x) \le \pi^{-1}(q_f) \le \pi^{-1}(r)$ implies $r \in Q_{\pi,x}$. Hence $r \in Q_{\pi,x} \subset Q_{\pi,q}$, which implies $d(\mathfrak{T}(Q_{\pi,q}), r) \ge d(\mathfrak{T}(Q_{\pi,x}), r)$. Also, $d(\mathfrak{T}(Q_{\pi,x}), r) \ge d_\pi(x)$ is from π being a D-ordering. Also, $d_\pi(x) = c_\pi(q_f)$ is from definition of x. Lastly, definition of j_L in Eq. (7) and $\pi^{-1}(q) < j_L$ implies $q \notin M_{\min}$, hence $c_\pi(q_f) \ge d_\pi(q) + \delta$ holds. From these, $d(\mathfrak{T}'(Q_{\pi',q}), r) \ge d_\pi(q)$ holds for $r \in C$ case. For $r \in Q_{\pi',q} - C$ case, we show as $d(\mathfrak{T}'(Q_{\pi',q}), r) = d(\mathfrak{T}(Q_{\pi,q}), r) \ge d(\mathfrak{T}(Q_{\pi,q}), q)$. Since r is not in C and from Eq. (23), $d(\mathfrak{T}'(Q_{\pi',q}), r) = d(\mathfrak{T}(Q_{\pi,q}), r)$ holds. Then from π being a D-ordering, $d(\mathfrak{T}(Q_{\pi,q}), r) \ge d_\pi(q)$ holds for $r \in Q_{\pi',q} - C$ case. Hence for either case, $d(\mathfrak{T}'(Q_{\pi',q}), r) \ge d_\pi(q)$ holds. This with $d_\pi(q) = d(\mathfrak{T}'(Q_{\pi',q}), q)$ implies $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$ for $\pi^{-1}(q) < j_L$.

(ii) For the case of $j_L \le \pi^{-1}(q) \le j_H$, we show Eq. (22) by $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q} \cap R} d(\mathfrak{T}'(Q_{\pi',q}), r)$ and $\min_{r \in Q_{\pi',q} - R} d(\mathfrak{T}'(Q_{\pi',q}), r) \ge c_\pi(q_f) \ge d(\mathfrak{T}'(Q_{\pi',q}), q)$. At first, Lemma B.1 (iii) implies $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi',q} \cap R} d(\mathfrak{T}'(Q_{\pi',q}), r)$. Also, since $M_{\max} = M(c_\pi(q_f))$, Lemma B.1 (iv) implies $\min_{r \in Q_{\pi',q} - R} d(\mathfrak{T}'(Q_{\pi',q}), r) \ge c_\pi(q_f)$. For $d(\mathfrak{T}'(Q_{\pi',q}), q) \le c_\pi(q_f)$, let $r := \arg\min_{q' \in Q_{\pi',q}} \pi^{-1}(q')$ be the slice index earliest in π among $Q_{\pi',q}$. We show $d(\mathfrak{T}'(Q_{\pi',q}), q) \le d(\mathfrak{T}'(Q_{\pi',q}), r) \le d(\mathfrak{T}(Q_{\pi',q}), r) \le d_\pi(r) \le c_\pi(q_f)$. $d(\mathfrak{T}'(Q_{\pi',q}), q) \le d(\mathfrak{T}'(Q_{\pi',q}), r)$ is from Eq. (16) in Lemma B.1 (v). And slice sums either decrease or remain the same under $((i_1, ..., i_N), \delta, -)$, so $d(\mathfrak{T}'(Q_{\pi',q}), r) \le d(\mathfrak{T}(Q_{\pi',q}), r)$ holds. And Eq. (17) in Lemma B.1 (v) implies $d(\mathfrak{T}(Q_{\pi',q}), r) \le d_\pi(r)$. Then, π and π' coinciding on $[|Q|] \backslash [j_L, j_H]$, $\pi^{-1}(r) \le \pi^{-1}(q) \le j_H$ holds. Then Eq. (8) implies that $r \notin M_{\max}$. Hence $\pi^{-1}(r) \le \pi^{-1}(q_f)$ or $d_\pi(r) < c_\pi(q_f)$ holds. For $\pi^{-1}(r) \le \pi^{-1}(q_f)$ case, definition of $c_\pi(\cdot)$ implies that $d_\pi(r) \le c_\pi(r) \le c_\pi(q_f)$, hence in any case $d_\pi(r) \le c_\pi(q_f)$ holds. Hence for for $j_L \le \pi^{-1}(q) \le j_H$ we have

$$d(\mathfrak{T}'(Q_{\pi',q}), q) \le c_\pi(q_f). \tag{24}$$

Then Lemma B.1 (iii), (iv), and Eq. (24) imply that $d(\mathfrak{T}'(Q_{\pi',q}), q) = \min_{r \in Q_{\pi,q}} d(\mathfrak{T}'(Q_{\pi',q}), r)$ holds for $j_L \le \pi^{-1}(q) \le j_H$.

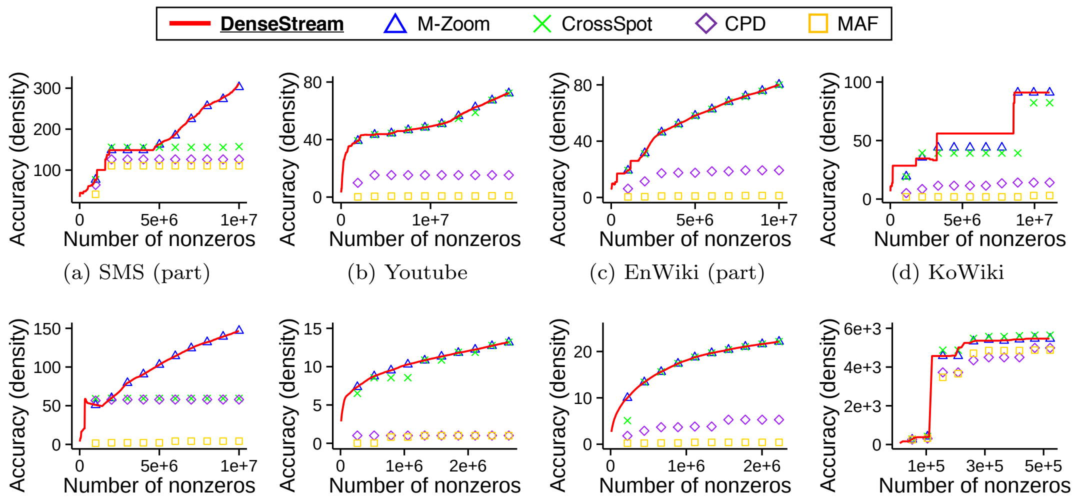
(iii) For the case of $\pi^{-1}(q) > j_H$, Lemma B.1 (vi) implies Eq. (19). ∎

C. ADDITIONAL EXPERIMENTS

We design additional experiments to answer the following question:

— **Q5. Accuracy in the Case of Decrement**: Does DENSESTREAM accurately maintain a dense subtensor when the values of tensor entries decrease?

Experimental settings were the same as in the main paper.



(a) SMS (part) (b) Youtube (c) EnWiki (part) (d) KoWiki

(e) YahooM. (part) (f) Android (g) Yelp (h) TCP

Fig. 6: **DenseStream is 'any-time' and accurate in the case of decrement.** While tensors change, DenseStream maintains and instantly updates a dense subtensor, whereas batch methods update dense subtensors once in a time interval. Moreover, while the values of tensor entries decrease (and eventually become zero), the subtensors maintained by DenseStream have density (red lines) similar to the density (points) of the subtensors found by the best batch methods. In the main paper, we show that DenseStream accurately update dense subtensors while the values of tensor entries increase.

C.1. Q5. Accuracy in the Case of Decrement

We tracked the density of the dense subtensor maintained by DenseStream while the values of tensor entries decrease (and eventually become zero), and compared it with the densities of the dense subtensors found by batch algorithms [Jiang et al. 2015; Maruhashi et al. 2011; Shin et al. 2016] As seen in Figure 6, the subtensors maintained by DenseStream have density (red lines) similar to the density (points) of the subtensors found by the best batch algorithms. Moreover, DenseStream is 'any time'. That is, as seen in Figure 6, DenseStream updates the dense subtensor instantly, while the batch algorithms cannot update their dense subtensors until the next batch processes end. In the main paper, we show that DenseStream also accurately updates dense subtensors while the values of tensor entries increase.

REFERENCES

Meng Jiang, Alex Beutel, Peng Cui, Bryan Hooi, Shiqiang Yang, and Christos Faloutsos. 2015. A general suspiciousness metric for dense blocks in multimodal data. In *ICDM*.

Koji Maruhashi, Fan Guo, and Christos Faloutsos. 2011. Multiaspectforensics: Pattern mining on large-scale heterogeneous networks with tensor analysis. In *ASONAM*.

Kijung Shin, Bryan Hooi, and Christos Faloutsos. 2016. M-Zoom: Fast Dense-Block Detection in Tensors with Quality Guarantees. In *ECML/PKDD*.

Kijung Shin, Bryan Hooi, Jisu Kim, and Christos Faloutsos. 2017. DenseAlert: Incremental Dense-Subtensor Detection in Tensor Stream. In *KDD*.