Think before You Discard: Accurate Triangle Counting in Graph Streams with Deletions (Supplementary Document)

Kijung Shin[1(✉)], Jisu Kim[2], Bryan Hooi[2], and Christos Faloutsos[1]

School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA
[1]{kijungs,christos}@cs.cmu.edu, [2]{jisuk1,bhooi}@andrew.cmu.edu

Abstract. In this supplementary document, we provide proofs, a variance analysis, and additional experimental results, all of which supplement the main paper [6].

A Proofs

For each variable (e.g., $\bar{c}$) in Algorithms 1 and 2, we use superscript (t) (e.g., $\bar{c}^{(t)}$) to denote the value of the variable after the t-th element $e^{(t)}$ is processed. For any time $t \geq 1$, let $X^{(t)}$ be the random number in $Bernoulli(r)$ drawn in line 11 of Algorithm 1 while the t-th element $e^{(t)}$ is processed, and for each edge $\{u, v\}$, let $l_{uv}^{(t)}$ be the last time that $\{u, v\}$ is added to or removed from $\mathcal{G}$ at time t or earlier. That is,

$$l_{uv}^{(t)} := \max(\{1 \leq s \leq t : e^{(s)} = (\{u, v\}, +) \text{ or } e^{(s)} = (\{u, v\}, -)\}). \tag{10}$$

Lemma 4. *In Algorithm 1, for each time $t \geq 1$ and any edge $\{u, v\} \in \mathcal{E}^{(t)}$, $\{u, v\} \in \mathcal{S}^{(t)}$ if and only if $X^{(l_{uv}^{(t)})} = 1$. That is,*

$$\{u, v\} \in \mathcal{S}^{(t)} \iff X^{(l_{uv}^{(t)})} = 1, \ \forall t \geq 1, \ \forall \{u, v\} \in \mathcal{E}^{(t)} \tag{11}$$

Proof. Note that $\{u, v\} \in \mathcal{E}^{(t)}$ implies that $e^{(l_{uv}^{(t)})} = (\{u, v\}, +)$, i.e. the edge $\{u, v\}$ is added at time $l_{uv}^{(t)}$. Then $\{u, v\} \notin \mathcal{E}^{(l_{uv}^{(t)}-1)}$, and since $\mathcal{S}^{(s)} \subset \mathcal{E}^{(s)}$ for all $s \geq 1$, $\{u, v\} \notin \mathcal{S}^{(l_{uv}^{(t)}-1)}$ as well. Therefore,

$$\{u, v\} \in \mathcal{S}^{(l_{uv}^{(t)})} \iff X^{(l_{uv}^{(t)})} = 1. \tag{12}$$

Also, from Eq. (10), $e^{(s)} \neq (\{u, v\}, \delta)$ if $l_{uv}^{(t)} < s \leq t$, and hence $\{u, v\}$ is not added after time $l_{uv}^{(t)}$ in Algorithm 1. Hence for all $s \in [l_{uv}^{(t)}, t]$,

$$\{u, v\} \in \mathcal{S}^{(s)} \iff \{u, v\} \in \mathcal{S}^{(l_{uv}^{(t)})}. \tag{13}$$

Combining Eq. (12) and Eq. (13) with $s = t$ gives Eq. (11). ■

A.1 Proof of Lemma 1

Proof. Applying Lemma 4 to the edges $\{u, v\}$ and $\{w, x\}$ gives

$$\{u,v\} \in \mathcal{S}^{(t)} \iff X^{(l_{uv}^{(t)})} = 1 \quad \text{and} \quad \{w,x\} \in \mathcal{S}^{(t)} \iff X^{(l_{wx}^{(t)})} = 1. \tag{14}$$

Then, since X_s's are independent $Bernoulli(r)$ and $l_{uv}^{(t)} \neq l_{wx}^{(t)}$, applying Eq. (14) with independence of $X^{(l_{uv}^{(t)})}$ and $X^{(l_{wx}^{(t)})}$ gives

$$\begin{aligned} Pr\left[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}\right] &= Pr\left[X^{(l_{uv}^{(t)})} = 1 \cap X^{(l_{wx}^{(t)})} = 1\right] \\ &= Pr\left[X^{(l_{uv}^{(t)})} = 1\right] Pr\left[X^{(l_{wx}^{(t)})} = 1\right] \quad = r^2. \end{aligned}$$

■

A.2 Proof of Lemma 2

As in the main paper, for each time $t \geq 1$, let $\mathcal{E}^{(t)}$ be the set of edges remaining (without being deleted) in the input graph stream and $\mathcal{S}^{(t)} \subset \mathcal{E}^{(t)}$ be the set of samples maintained by Algorithm 2 after the t-th element is processed. Also let $y^{(t)} = \min(k, |\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)})$.

Lemma 5 (Properties in Random Pairing [1]). *In Algorithm 2, the expected value and variance of the size of the samples are as follows:*

$$\mathbb{E}[|\mathcal{S}^{(t)}|] = \frac{|\mathcal{E}^{(t)}| \cdot y^{(t)}}{|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)}}, \ \forall t \geq 1. \tag{15}$$

$$Var[|\mathcal{S}^{(t)}|] = \frac{(n_b^{(t)} + n_g^{(t)}) \cdot y^{(t)} \cdot (|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)} - y^{(t)}) \cdot |\mathcal{E}^{(t)}|}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)})^2 \cdot (|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)} - 1)}, \forall t \geq 1. \tag{16}$$

At each fixed time, all equal-sized subsets of the remaining elements in the input graph stream have the same probability to be the set of samples maintained in Algorithm 2. Formally,

$$Pr[\mathcal{S}^{(t)} = \mathcal{A}] = Pr[\mathcal{S}^{(t)} = \mathcal{B}], \ \forall t \geq 1, \ \forall \mathcal{A} \neq \mathcal{B} \subset \mathcal{E}^{(t)} \ s.t. \ |\mathcal{A}| = |\mathcal{B}|. \tag{17}$$

Lemma 6 (Uniformity in Random Pairing). *At each fixed time, all equal-sized subsets of the remaining elements in the input graph stream have the same probability to be a subset of the samples maintained in Algorithm 2. Formally,*

$$Pr[\mathcal{A} \subset \mathcal{S}^{(t)}] = Pr[\mathcal{B} \subset \mathcal{S}^{(t)}], \ \forall t \geq 1, \ \forall \mathcal{A} \neq \mathcal{B} \subset \mathcal{E}^{(t)} \ s.t. \ |\mathcal{A}| = |\mathcal{B}|. \tag{18}$$

Proof. Let $e_i^{\mathcal{A}}$ be the family of size-i subsets of $\mathcal{E}^{(t)}$ including $\mathcal{A}$, and let $e_i^{\mathcal{B}}$ be the family of size-i subsets of $\mathcal{E}^{(t)}$ including $\mathcal{B}$. Then, Eq. (18) is obtained as follows:

$$\begin{aligned} Pr[\mathcal{A} \subset \mathcal{S}^{(t)}] &= \sum_i \sum_{\mathcal{C} \in e_i^{\mathcal{A}}} Pr[\mathcal{C} = \mathcal{S}^{(t)}] \\ &= \sum_i \sum_{\mathcal{C} \in e_i^{\mathcal{B}}} Pr[\mathcal{C} = \mathcal{S}^{(t)}] = Pr[\mathcal{B} \subset \mathcal{S}^{(t)}], \end{aligned}$$

where the second equality is from Eq. (17) and $|e_i^{\mathcal{A}}| = |e_i^{\mathcal{B}}|$. ■

Lemma 7 (Sampling Probability of Each Edge). *The probability that each edge is sampled in Algorithm 2 is as follows:*

$$Pr[\{u,v\} \in \mathcal{S}^{(t)}] = \frac{y^{(t)}}{|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)}},\ \forall t \geq 1,\ \forall \{u,v\} \in \mathcal{E}^{(t)}. \tag{19}$$

Proof. Let $\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)})$ be a random variable which is 1 if $\{u,v\} \in \mathcal{S}^{(t)}$ and 0 otherwise. By definition,

$$|\mathcal{S}^{(t)}| = \sum_{\{u,v\} \in \mathcal{E}^{(t)}} \mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)}). \tag{20}$$

Then, by linearity of expectation and Eq. (20),

$$\mathbb{E}[|\mathcal{S}^{(t)}|] = \sum_{\{u,v\} \in \mathcal{E}^{(t)}} \mathbb{E}[\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)})] = \sum_{\{u,v\} \in \mathcal{E}^{(t)}} Pr[\{u,v\} \in \mathcal{S}^{(t)}]. \tag{21}$$

Then, Eq. (19) is obtained as follows:

$$\begin{aligned} Pr[\{u,v\} \in \mathcal{S}^{(t)}] &= \frac{1}{|\mathcal{E}^{(t)}|} \sum_{\{w,x\} \in \mathcal{E}^{(t)}} Pr[\{w,x\} \in \mathcal{S}^{(t)}] \\ &= \frac{\mathbb{E}[|\mathcal{S}^{(t)}|]}{|\mathcal{E}^{(t)}|} = \frac{y^{(t)}}{|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)}}, \end{aligned}$$

where the first, second, and last equalities are from Eq. (18), Eq. (21), and Eq. (15), respectively. ■

Proof of Lemma 2:

Proof. Let $\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)})$ be a random variable which is 1 if $\{u,v\} \in \mathcal{S}^{(t)}$ and 0 otherwise. For calculating $Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}]$, we expand the covariance sum $\sum_{\{u,v\} \neq \{w,x\}} Cov(\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)}), \mathbf{1}(\{w,x\} \in \mathcal{S}^{(t)}))$ in two ways and compare them.

First, we use the expansion of the variance of $\mathcal{S}^{(t)}$. From Eq. (20),

$$\begin{aligned} Var[|\mathcal{S}^{(t)}|] = &\sum_{\{u,v\} \in \mathcal{E}^{(t)}} Var[\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)})] \\ &+ \sum_{\{u,v\} \neq \{w,x\}} Cov(\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)}), \mathbf{1}(\{w,x\} \in \mathcal{S}^{(t)})), \end{aligned}$$

and hence the covariance sum can be expanded as

$$\begin{aligned} &\sum_{\{u,v\} \neq \{w,x\}} Cov(\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)}), \mathbf{1}(\{w,x\} \in \mathcal{S}^{(t)})) \\ &= Var[|\mathcal{S}^{(t)}|] - \sum_{\{u,v\} \in \mathcal{E}^{(t)}} Var[\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)})]. \end{aligned} \tag{22}$$

For the second term of Eq. (22), $Var[x] = \mathbb{E}[x^2] - (\mathbb{E}[x])^2$ implies

$$Var[\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)})] = Pr[\{u,v\} \in \mathcal{S}^{(t)}] - Pr[\{u,v\} \in \mathcal{S}^{(t)}]^2. \tag{23}$$

Hence applying Eq. (19) and Eq. (23) to Eq. (22) gives the covariance sum as

$$\begin{aligned}
&\sum_{\{u,v\} \neq \{w,x\}} Cov(\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)}), \mathbf{1}(\{w,x\} \in \mathcal{S}^{(t)})) \\
&= Var[|\mathcal{S}^{(t)}|] - \sum_{\{u,v\} \in \mathcal{E}^{(t)}} Var[\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)})] \\
&= Var[|\mathcal{S}^{(t)}|] - \sum_{\{u,v\} \in \mathcal{E}^{(t)}} \Big(Pr[\{u,v\} \in \mathcal{S}^{(t)}] - Pr[\{u,v\} \in \mathcal{S}^{(t)}]^2 \Big) \\
&= Var[|\mathcal{S}^{(t)}|] - |\mathcal{E}^{(t)}| \cdot \frac{y^{(t)} \cdot (|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)} - y^{(t)})}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)})^2}.
\end{aligned} \tag{24}$$

Second, we directly expand the covariance sum. Expanding the covariance sum with $Cov(x,y) = \mathbb{E}[xy] - \mathbb{E}[x] \cdot \mathbb{E}[y]$ and applying Eq. (19) give

$$\begin{aligned}
&\sum_{\{u,v\} \neq \{w,x\}} Cov(\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)}), \mathbf{1}(\{w,x\} \in \mathcal{S}^{(t)})) \\
&= \sum_{\{u,v\} \neq \{w,x\}} \Big(Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}] - Pr[\{u,v\} \in \mathcal{S}^{(t)}] \cdot Pr[\{w,x\} \in \mathcal{S}^{(t)}] \Big) \\
&= \sum_{\{u,v\} \neq \{w,x\}} \left(Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}] - \left(\frac{y^{(t)}}{|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)}} \right)^2 \right) \\
&= \sum_{\{u,v\} \neq \{w,x\}} \Big(Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}] \Big) - \frac{y^{(t)} \cdot y^{(t)} \cdot |\mathcal{E}^{(t)}| \cdot (|\mathcal{E}^{(t)}| - 1)}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)})^2}.
\end{aligned} \tag{25}$$

Now, the probability sum $\sum_{\{u,v\} \neq \{w,x\}} Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}]$ can be obtained by comparing two expansions Eq. (24) and Eq. (25) of the covariance sum and applying Eq. (16) as

$$\begin{aligned}
&\sum_{\{u,v\} \neq \{w,x\}} Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}] \\
&= \sum_{\{u,v\} \neq \{w,x\}} Cov(\mathbf{1}(\{u,v\} \in \mathcal{S}^{(t)}), \mathbf{1}(\{w,x\} \in \mathcal{S}^{(t)})) + \frac{y^{(t)} \cdot y^{(t)} \cdot |\mathcal{E}^{(t)}| \cdot (|\mathcal{E}^{(t)}| - 1)}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)})^2} \\
&= Var[|\mathcal{S}^{(t)}|] - |\mathcal{E}^{(t)}| \cdot \frac{y^{(t)} \cdot (|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)} - y^{(t)})}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)})^2} + \frac{y^{(t)} \cdot y^{(t)} \cdot |\mathcal{E}^{(t)}| \cdot (|\mathcal{E}^{(t)}| - 1)}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)})^2} \\
&= \frac{y^{(t)} \cdot (y^{(t)} - 1) \cdot |\mathcal{E}^{(t)}| \cdot (|\mathcal{E}^{(t)}| - 1)}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)}) \cdot (|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)} - 1)}.
\end{aligned} \tag{26}$$

Then, Eq. (3) is obtained by Eq. (26) as follows:

$$\begin{aligned}
&Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}] \\
&= \frac{1}{|\mathcal{E}^{(t)}| \cdot (|\mathcal{E}^{(t)}| - 1)} \left(\sum_{\{u,v\} \neq \{w,x\}} Pr[\{u,v\} \in \mathcal{S}^{(t)} \cap \{w,x\} \in \mathcal{S}^{(t)}] \right) \\
&= \frac{1}{|\mathcal{E}^{(t)}| \cdot (|\mathcal{E}^{(t)}| - 1)} \cdot \frac{y^{(t)} \cdot (y^{(t)} - 1) \cdot |\mathcal{E}^{(t)}| \cdot (|\mathcal{E}^{(t)}| - 1)}{(|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)}) \cdot (|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)} - 1)} \\
&= \frac{y^{(t)}}{|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)}} \cdot \frac{y^{(t)} - 1}{|\mathcal{E}^{(t)}| + n_b^{(t)} + n_g^{(t)} - 1} = p^{(t)},
\end{aligned}$$

where the first equality is from Eq. (18). ■

A.3 Proof of Lemma 3

Proof. When $t = 1$, then $\mathcal{T}^{(1)} = \mathcal{A}^{(1)} = \mathcal{D}^{(1)} = \emptyset$ holds, and hence Eq. (4) trivially holds. Hence we assume that $t \geq 2$ from now on. First, we show that for each time $s \geq 2$,

$$|\mathcal{T}^{(s)}| - |\mathcal{T}^{(s-1)}| = |\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)}| - |\mathcal{D}^{(s)} \backslash \mathcal{D}^{(s-1)}|. \tag{27}$$

To show this, we show the following relations,

$$|\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)}| = |\mathcal{T}^{(s)} \backslash \mathcal{T}^{(s-1)}|, \tag{28}$$

$$|\mathcal{D}^{(s)} \backslash \mathcal{D}^{(s-1)}| = |\mathcal{T}^{(s-1)} \backslash \mathcal{T}^{(s)}|. \tag{29}$$

For Eq. (28), note that

$$\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)} = \Big\{ (\{u,v,w\}, s) : \{u,v,w\} \notin \mathcal{T}^{(s-1)} \text{ and } \{u,v,w\} \in \mathcal{T}^{(s)} \Big\}$$

and

$$\mathcal{T}^{(s)} \backslash \mathcal{T}^{(s-1)} = \Big\{ \{u,v,w\} : \{u,v,w\} \notin \mathcal{T}^{(s-1)} \text{ and } \{u,v,w\} \in \mathcal{T}^{(s)} \Big\}$$

hold, and hence Eq. (28) holds. Similarly,

$$\mathcal{D}^{(s)} \backslash \mathcal{D}^{(s-1)} = \Big\{ (\{u,v,w\}, s) : \{u,v,w\} \in \mathcal{T}^{(s-1)} \text{ and } \{u,v,w\} \notin \mathcal{T}^{(s)} \Big\}$$

and

$$\mathcal{T}^{(s-1)} \backslash \mathcal{T}^{(s)} = \Big\{ \{u,v,w\} : \{u,v,w\} \in \mathcal{T}^{(s-1)} \text{ and } \{u,v,w\} \notin \mathcal{T}^{(s)} \Big\}$$

hold, and hence Eq. (29) holds. Then, Eq. (28) and Eq. (29) imply

$$\begin{aligned}
|\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)}| + |\mathcal{T}^{(s-1)}| &= |\mathcal{T}^{(s)} \backslash \mathcal{T}^{(s-1)}| + |\mathcal{T}^{(s-1)}| = |\mathcal{T}^{(s-1)} \cup \mathcal{T}^{(s)}| \\
&= |\mathcal{T}^{(s-1)} \backslash \mathcal{T}^{(s)}| + |\mathcal{T}^{(s)}| = |\mathcal{D}^{(s)} \backslash \mathcal{D}^{(s-1)}| + |\mathcal{T}^{(s)}|,
\end{aligned}$$

and hence Eq. (27) holds. Then, summing up Eq. (27) from $s = 2$ to t yields

$$|\mathcal{T}^{(t)}| - |\mathcal{T}^{(1)}| = \sum_{s=2}^{t} |\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)}| - \sum_{s=2}^{t} |\mathcal{D}^{(s)} \backslash \mathcal{D}^{(s-1)}|. \tag{30}$$

Then, $\left\{\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)}\right\}_{s=2}^{t}$ being disjoint over s implies

$$\sum_{s=2}^{t} |\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)}| = \left| \bigcup_{s=2}^{t} (\mathcal{A}^{(s)} \backslash \mathcal{A}^{(s-1)}) \right| = |\mathcal{A}^{(t)} \backslash \mathcal{A}^{(1)}|, \tag{31}$$

and similarly,

$$\sum_{s=2}^{t} |\mathcal{D}^{(s)} \backslash \mathcal{D}^{(s-1)}| = |\mathcal{D}^{(t)} \backslash \mathcal{D}^{(1)}|. \tag{32}$$

holds. Then, applying Eq. (31), Eq. (32), and $\mathcal{T}^{(1)} = \mathcal{A}^{(1)} = \mathcal{D}^{(1)} = \emptyset$ to Eq. (30) yields that for all $t \geq 2$,

$$|\mathcal{T}^{(t)}| = |\mathcal{A}^{(t)}| - |\mathcal{D}^{(t)}|,$$

which completes the proof of Eq. (4).

For Eq. (5), replacing $\mathcal{T}^{(s)}$ by $\mathcal{T}^{(s)}[u]$, $\mathcal{A}^{(s)}$ by $\mathcal{A}^{(s)}[u]$, and $\mathcal{D}^{(s)}$ by $\mathcal{D}^{(s)}[u]$ and repeating above give the proof. ■

B Variance Analysis

As in Sect. A, let $l_{uv}^{(t)}$ be the last time that edge $\{u, v\}$ is added to or removed from $\mathcal{G}$ at time t or earlier. And for each added or deleted triangle $(\{u, v, w\}, s) \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}$, we use $\mathbf{1}_{(\{u,v,w\},s)}$ to denote the time when its first edge has arrived and $\mathbf{2}_{(\{u,v,w\},s)}$ to denote the time when its second edge has arrived. Formally,

$$\mathbf{1}_{(\{u,v,w\},s)} := \min(l_{uv}^{(s)}, l_{vw}^{(s)}, l_{wu}^{(s)}), \quad \mathbf{2}_{(\{u,v,w\},s)} := \text{median}(l_{uv}^{(s)}, l_{vw}^{(s)}, l_{wu}^{(s)}).$$

Then, we define the type of each triangle pair in Definition 3.

Definition 3 (Types of Triangle Pairs). *The type of each ordered pair of two distinct triangles $\tau \neq \omega \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}$ is defined as follows:*

$$Type_{(\tau,\omega)} = \begin{cases} 1, & \text{if } \tau \in \mathcal{A}^{(t)} \text{ and } \omega \in \mathcal{A}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 1, \\ 2, & \text{if } \tau \in \mathcal{D}^{(t)} \text{ and } \omega \in \mathcal{D}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 1, \\ 3, & \text{if } \tau \in \mathcal{A}^{(t)} \text{ and } \omega \in \mathcal{D}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 1, \\ 4, & \text{if } \tau \in \mathcal{D}^{(t)} \text{ and } \omega \in \mathcal{A}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 1, \\ 5, & \text{if } \tau \in \mathcal{A}^{(t)} \text{ and } \omega \in \mathcal{A}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 2, \\ 6, & \text{if } \tau \in \mathcal{D}^{(t)} \text{ and } \omega \in \mathcal{D}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 2, \\ 7, & \text{if } \tau \in \mathcal{A}^{(t)} \text{ and } \omega \in \mathcal{D}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 2, \\ 8, & \text{if } \tau \in \mathcal{D}^{(t)} \text{ and } \omega \in \mathcal{A}^{(t)} \text{ and } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 2, \\ 9, & \text{otherwise (i.e., } |\{\mathbf{1}_\tau, \mathbf{2}_\tau\} \cap \{\mathbf{1}_\omega, \mathbf{2}_\omega\}| = 0). \end{cases} \tag{33}$$

Theorem 5 (Variance of ThinkD$_{\mathbf{fast}}$). *Let $n_i^{(t)}$ be the number of Type-i triangle pairs in $\mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}$. Likewise, Let $n_i^{(t)}[u]$ be the number of Type-i triangle pairs in $\mathcal{A}^{(t)}[u] \cup \mathcal{D}^{(t)}[u]$. Then,*

$$\begin{aligned} Var[\bar{c}^{(t)}] = &(|\mathcal{A}^{(t)}| + |\mathcal{D}^{(t)}|) \cdot \frac{1-r^2}{r^2} \\ &+ (n_1^{(t)} + n_2^{(t)} - n_3^{(t)} - n_4^{(t)}) \cdot \frac{1-r}{r} \\ &+ (n_5^{(t)} + n_6^{(t)} - n_7^{(t)} - n_8^{(t)}) \cdot \frac{1-r^2}{r^2}, \ \forall t \geq 1. \end{aligned} \tag{34}$$

Likewise,

$$\begin{aligned} Var[c^{(t)}[u]] = &(|\mathcal{A}^{(t)}[u]| + |\mathcal{D}^{(t)}[u]|) \cdot \frac{1-r^2}{r^2} \\ &+ (n_1^{(t)}[u] + n_2^{(t)}[u] - n_3^{(t)}[u] - n_4^{(t)}[u]) \cdot \frac{1-r}{r} \\ &+ (n_5^{(t)}[u] + n_6^{(t)}[u] - n_7^{(t)}[u] - n_8^{(t)}[u]) \cdot \frac{1-r^2}{r^2}, \ \forall t \geq 1, \ \forall u \in \mathcal{V}^{(t)}. \end{aligned} \tag{35}$$

Proof. As in Sect. A, for each time $t \geq 1$, let $X^{(t)}$ be the random number in $Bernoulli(r)$ drawn in line 11 of Algorithm 1 while the t-th element $e^{(t)}$ is processed. Then, from Lemma 4,

$$\{u, v\} \in \mathcal{S}^{(t)} \iff X^{(l_{uv}^{(t)})} = 1. \tag{36}$$

Now, for each $\tau = (\{u, v, w\}, s) \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}$, let γ_τ be the amount of change in each of $\bar{c}$, $c[u]$, $c[v]$, and $c[w]$ due to the discovery of τ in line 8 or line 9 of Algorithm 1. Let $\delta_\tau = +1$ when $\tau \in \mathcal{A}^{(t)}$, i.e. when the last edge is added, and let $\delta_\tau = -1$ when $\tau \in \mathcal{D}^{(t)}$, i.e. when the last edge is deleted. Let $\{u, v\}$ be the edge added or deleted at time s without loss of generality. Then, $\gamma_\tau = \frac{\delta_\tau}{r^2}$ if both $\{v, w\}, \{w, u\} \in \mathcal{S}^{(s)}$ and 0 otherwise. Hence, combined with Eq. (36),

$$\gamma_\tau = \frac{\delta_\tau}{r^2} X_{\mathbf{1}_\tau} X_{\mathbf{2}_\tau}. \tag{37}$$

Then from the definitions of γ_τ, $\bar{c}^{(t)} = \sum_{\tau \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} \gamma_\tau$, and its variance is

$$Var[\bar{c}^{(t)}] = \sum_{\tau \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} Var[\gamma_\tau] + \sum_{\tau \neq \omega \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} Cov[\gamma_\tau, \gamma_\omega]. \tag{38}$$

For the variance term in Eq. (38), note first that applying that $X_{\mathbf{1}_\tau}$ and $X_{\mathbf{2}_\tau}$ are independent $Bernoulli(r)$ to Eq. (37) gives $\mathbb{E}[\gamma_\tau]$ as

$$\mathbb{E}[\gamma_\tau] = \mathbb{E}\left[\frac{\delta_\tau}{r^2} X_{\mathbf{1}_\tau} X_{\mathbf{2}_\tau}\right] = \frac{\delta_\tau}{r^2} \mathbb{E}\left[X_{\mathbf{1}_\tau}\right] \mathbb{E}\left[X_{\mathbf{2}_\tau}\right] = \delta_\tau. \tag{39}$$

Then, further applying $\delta_\tau^2 = 1$ and $X_s^2 = X_s$ to Eq. (37) and again applying that $X_{\mathbf{1}_\tau}$ and $X_{\mathbf{2}_\tau}$ are independent $Bernoulli(r)$ give $Var[\gamma_\tau]$ as

$$\begin{aligned} Var[\gamma_\tau] = \mathbb{E}[\gamma_\tau^2] - (\mathbb{E}[\gamma_\tau])^2 &= \mathbb{E}\left[\frac{\delta_\tau^2}{r^4} X_{\mathbf{1}_\tau} X_{\mathbf{2}_\tau}\right] - \delta_\tau^2 = \frac{1}{r^4}\mathbb{E}\left[X_{\mathbf{1}_\tau}\right]\mathbb{E}\left[X_{\mathbf{2}_\tau}\right] - 1 \\ &= \frac{1-r^2}{r^2}. \end{aligned}$$

Hence the variance term in Eq. (38) is computed as

$$\sum_{\tau \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} Var[\gamma_\tau] = \sum_{\tau \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} \frac{1-r^2}{r^2} = (|\mathcal{A}^{(t)}| + |\mathcal{D}^{(t)}|) \cdot \frac{1-r^2}{r^2}. \tag{40}$$

For the covariance term in Eq. (38), applying Eq. (37) and Eq. (39) and using the fact that all the X_s's are independent and identically distributed as $Bernoulli(r)$ and $X_s^2 = X_s$ yield the $Cov[\gamma_\tau, \gamma_\omega]$ as

$$\begin{aligned} Cov[\gamma_\tau, \gamma_\omega] = \mathbb{E}[\gamma_\tau \gamma_\omega] - \mathbb{E}[\gamma_\tau]\mathbb{E}[\gamma_\omega] &= \mathbb{E}\left[\frac{\delta_\tau \delta_\omega}{r^4} X_{\mathbf{1}_\tau} X_{\mathbf{2}_\tau} X_{\mathbf{1}_\omega} X_{\mathbf{2}_\omega}\right] - \delta_\tau \delta_\omega \\ &= \delta_\tau \delta_\omega \left(\frac{1}{r^4}\mathbb{E}\left[\prod_{i \in \{\mathbf{1}_\gamma, \mathbf{2}_\gamma\} \cup \{\mathbf{1}_\omega, \mathbf{2}_\omega\}} X_i\right] - 1\right) \\ &= \delta_\tau \delta_\omega \left(\frac{1}{r^4} \prod_{i \in \{\mathbf{1}_\gamma, \mathbf{2}_\gamma\} \cup \{\mathbf{1}_\omega, \mathbf{2}_\omega\}} \mathbb{E}\left[X_i\right] - 1\right) = \delta_\tau \delta_\omega \left(\frac{r^{|\{\mathbf{1}_\gamma, \mathbf{2}_\gamma\} \cup \{\mathbf{1}_\omega, \mathbf{2}_\omega\}|}}{r^4} - 1\right). \end{aligned}$$

Then $\delta_\tau \delta_\omega = 1$ if $\tau, \omega \in \mathcal{A}^{(t)}$ or $\tau, \omega \in \mathcal{D}^{(t)}$, and $\delta_\tau \delta_\omega = -1$ if $\tau \in \mathcal{A}^{(t)}$, $\omega \in \mathcal{D}^{(t)}$ or $\tau \in \mathcal{D}^{(t)}$, $\omega \in \mathcal{A}^{(t)}$. Hence $Cov[\gamma_\tau, \gamma_\omega]$ can be calculated as

$$Cov[\gamma_\tau, \gamma_\omega] = \begin{cases} \frac{1-r}{r}, & \text{if } Type_{(\tau,\omega)} = 1 \text{ or } 2, \\ -\frac{1-r}{r}, & \text{if } Type_{(\tau,\omega)} = 3 \text{ or } 4, \\ \frac{1-r^2}{r^2}, & \text{if } Type_{(\tau,\omega)} = 5 \text{ or } 6, \\ -\frac{1-r^2}{r^2}, & \text{if } Type_{(\tau,\omega)} = 7 \text{ or } 8, \\ 0, & \text{if } Type_{(\tau,\omega)} = 9. \end{cases}$$

Hence the covariance term in Eq. (38) is computed as

$$\begin{aligned} \sum_{\tau \neq \omega \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} Cov[\gamma_\tau, \gamma_\omega] =& (n_1^{(t)} + n_2^{(t)} - n_3^{(t)} - n_4^{(t)}) \cdot \frac{1-r}{r} \\ &+ (n_5^{(t)} + n_6^{(t)} - n_7^{(t)} - n_8^{(t)}) \cdot \frac{1-r^2}{r^2}. \end{aligned} \tag{41}$$

Hence applying Eq. (40) and Eq. (41) to Eq. (38) gives

$$\begin{aligned} Var[\bar{c}^{(t)}] = & \sum_{\tau \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} Var[\gamma_\tau] + \sum_{\tau \neq \omega \in \mathcal{A}^{(t)} \cup \mathcal{D}^{(t)}} Cov[\gamma_\tau, \gamma_\omega] \\ = & (|\mathcal{A}^{(t)}| + |\mathcal{D}^{(t)}|) \cdot \frac{1-r^2}{r^2} + (n_1^{(t)} + n_2^{(t)} - n_3^{(t)} - n_4^{(t)}) \cdot \frac{1-r}{r} \\ & + (n_5^{(t)} + n_6^{(t)} - n_7^{(t)} - n_8^{(t)}) \cdot \frac{1-r^2}{r^2}, \end{aligned}$$

which completes the proof of Eq. (34).

For Eq. (35), replacing $\bar{c}^{(t)}$ by $c^{(t)}[u]$, $\mathcal{A}^{(t)}$ by $\mathcal{A}^{(t)}[u]$, and $\mathcal{D}^{(s)}$ by $\mathcal{D}^{(s)}[u]$ and repeating above give the proof. ■

C Additional Experimental Results

Figures 5, 6, and 7 show additional experimental results that supplement Sect. 5.3, Sect. 5.4, and Sect. 5.5, respectively, of the main paper.

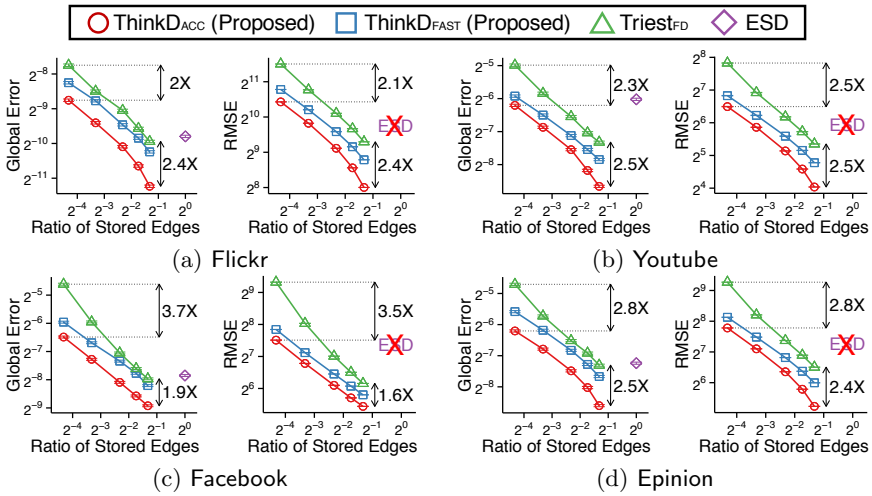
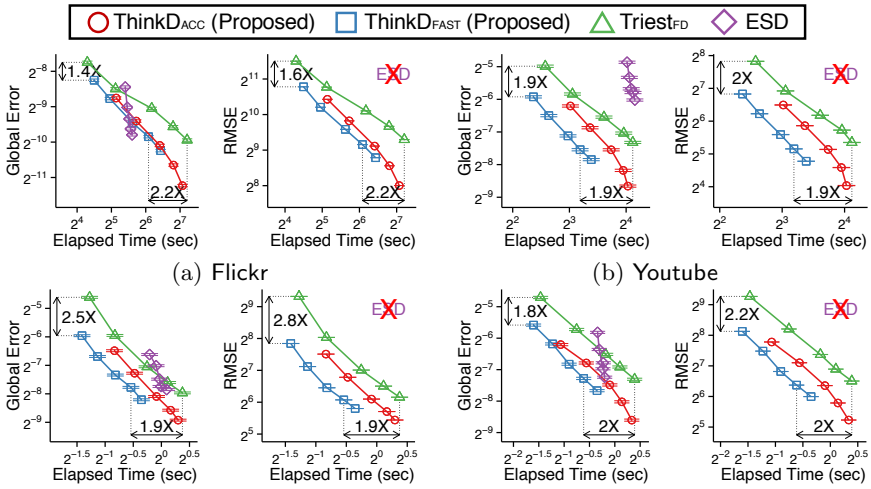
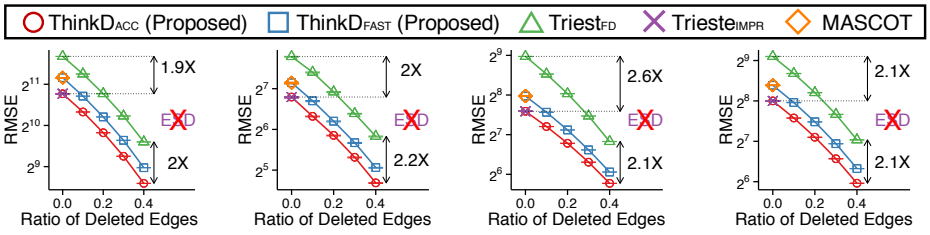


Fig. 5: **ThinkD is accurate**. ThinkD gives the best trade-off between space and accuracy. In particular, ThinkD$_{\text{ACC}}$ is up to **4.3× more accurate** than Triest$_{\text{FD}}$ within the same memory budget. Error bars denote ±1 standard error. ESD is inapplicable to local triangle counting.



(a) Flickr (b) Youtube (c) Facebook (d) Epinion

Fig. 6: **ThinkD is fast**. THINKD gives the best trade-off between speed and accuracy. In particular, $\text{THINKD}_{\text{FAST}}$ is up to **2.2× faster** than $\text{TRIEST}_{\text{FD}}$ when they are similarly accurate. Error bars denote ±1 standard error. ESD is inapplicable to local triangle counting.



(a) Flickr (b) Youtube (c) Facebook (d) Epinion

Fig. 7: **ThinkD is consistently accurate regardless of the ratio of deleted edges.** Error bars denote ±1 standard error. $\text{TRIEST}_{\text{IMPR}}$ and MASCOT are inapplicable when there are deletions. ESD is inapplicable to local triangle counting.

D Detecting Sudden Emergence of Dense Subgraphs

We describe how ThinkD can be used to the task of spotting the sudden emergence of dense subgraphs, which indicate a variety of anomalies [2,3,4,5]. Our method exploits the fact that dense subgraphs contain many triangles.

Consider a graph whose edges have timestamps, and for each edge e, let t_e be its timestamp. We create a fully dynamic graph stream with deletions as follows:

(1) For each edge e, create an addition $(e, +)$ with timestamp t_e,
(2) For each edge e, create a deletion $(e, -)$ with timestamp $t_e + \Delta T$,
(3) Sort the additions and deletions from the previous steps by their timestamps.

Then, we process the created stream using ThinkD (either $\text{ThinkD}_{\text{FAST}}$ or $\text{ThinkD}_{\text{ACC}}$). This makes ThinkD maintain the estimated number of triangles created within ΔT time units. We track the changes in the maintained estimate of the global triangle count and report the spikes. Although we describe our method in an offline setting, we can easily extend it to online settings where new edges are added to the input graph while we process it. In online settings, for each new edge, we process its insertion when it arrives, while we process its deletion after ΔT time units.

For empirical evaluation of our method, we measured how rapidly and accurately our method detect anomalies, compared to its competitors: DenseAlert [5], M-Biz [3], D-Cube [4], and CrossSpot [2], all of which are designed for detecting dense subgraphs or more generally dense subtensors created within a short time.[1] Specifically, we randomly injected 10 cliques of sizes from 6 to 15 into Facebook and Epinion datasets,[2] which have timestamps, so that each clique is formed (from scratch) within a day. Then, we measured the running time of each method and the precision at the top-10 outputs obtained by each method (i.e., the ratio of outputs that correspond to the injected cliques).

As summarized in Fig. 8(a), our method was the fastest and the most accurate in both datasets. Especially, our method was $\mathbf{8.3 - 19.6\times}$ **faster** than DenseAlert, which was the second most accurate method. Figs. 8(b)-(e) show that the spikes obtained by our method indicated the injections more accurately than the spikes obtained by DenseAlert.

References

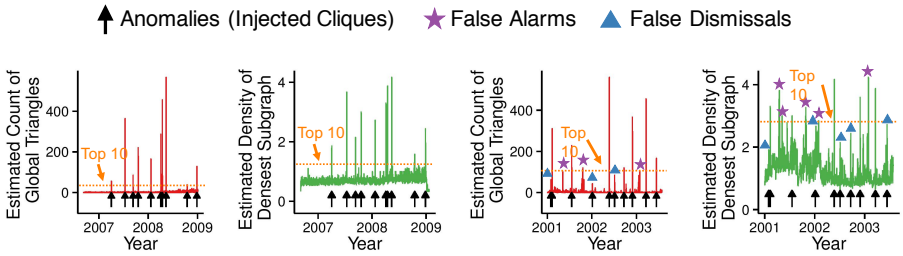
1. Gemulla, R., Lehner, W., Haas, P.J.: Maintaining bounded-size sample synopses of evolving datasets. The VLDB Journal 17(2), 173–201 (2008)

[1] We used the same machine used in the main paper, and we implemented all methods commonly in Java. We used $\text{ThinkD}_{\text{ACC}}$ with $k = 2000$ for triangle counting in our method, and we set ΔT to a day for both our method and DenseAlert. We used the geometric average degree as the density measure in M-Biz and D-Cube, since the measure led to the highest accuracy. Using each of M-Biz, D-Cube, and CrossSpot, we detected 10 dense subgraphs.

[2] We ignored the edges whose timestamps are missing.

Datasets	Facebook		Epinion	
Measures	Running Time (milliseconds)	Precision @ Top-10	Running Time (milliseconds)	Precision @ Top-10
CrossSpot [2]	54,975	0.00	11,173	0.00
D-Cube [4]	1,443	0.33	674	0.00
M-Biz [3]	1,298	0.31	667	0.41
DenseAlert [5]	12,080	**1.00**	1,843	0.59
Proposed	**617**	**1.00**	**223**	**0.75**

(a) Running Time and Accuracy Averaged over 10 Runs.



(b) Proposed (Facebook) (c) DenseAlert (Facebook) (d) Proposed (Epinion) (e) DenseAlert (Epinion)

Fig. 8: **Our anomaly-detection method based on ThinkD is fast and accurate**, as summarized in (a). In (b)-(e), the arrows indicate the timestamps where sudden dense subgraphs are injected, and the stars indicate the false alarms in top-10 spikes. The triangles indicate false dismissals. The spikes obtained by our method indicate the injections more accurately than the spikes obtained by DenseAlert.

2. Jiang, M., Beutel, A., Cui, P., Hooi, B., Yang, S., Faloutsos, C.: A general suspiciousness metric for dense blocks in multimodal data. In: ICDM (2015)
3. Shin, K., Hooi, B., Faloutsos, C.: Fast, accurate, and flexible algorithms for dense subtensor mining. TKDD 12(3), 28:1–28:30 (2018)
4. Shin, K., Hooi, B., Kim, J., Faloutsos, C.: D-cube: Dense-block detection in terabyte-scale tensors. In: WSDM (2017)
5. Shin, K., Hooi, B., Kim, J., Faloutsos, C.: Densealert: Incremental dense-subtensor detection in tensor streams. In: KDD (2017)
6. Shin, K., Kim, J., Hooi, B., Faloutsos, C.: Think Before You Discard: Accurate Triangle Counting in Graph Streams with Deletions. In: ECML/PKDD (2018)