

## [Issue #16](#) (open): [REVIEW] Gatherplot Review

[@jov-anonymous-reviewer-AAAA](#) on Oct 16, 2023 07:57:

[opened]

[@jov-anonymous-reviewer-AAAA](#) on Oct 16, 2023 07:57:

---

### Conflicts of interest

- I declare that I have no known conflicts of interest with the authors.

---

### Reviewed version

0ee8b81

---

### Review

#### Short Summary

This paper presents Gatherplot, a variant of conventional scatterplot that alleviates the overplotting problem. By “gathering” and “packing” the data points in the scatterplot, users can more easily and accurately examine grouped objects. The Gatherplot system was evaluated through a crowdsourced experiment conducted on the Amazon Mturk Platform, where the results indicate Gatherplot’s superiority against the jittering solution. The paper ends by introducing Gatherlines, a focus+context solution utilizing Gathering.

#### Verdict

This paper presents an interesting visualization technique and the technique is rigorously validated through a large-scale crowdsourced user study. However, the paper currently holds several limitations (see major and minor weaknesses), and there exist questions on the novelty of the proposed system. Therefore, the paper currently needs substantial revision before publishing. Thus, the final decision is **Major revision**.

#### Paper Strength

**(S1) Well-designed experiment** The paper includes rigorous experiments using the Amazon Mechanical Turk platform. Due to its crowdsourced nature, the study results seem to be more credible. Despite there being some missing points in the study design (Check Weakness), the study is overall well-designed and executed.

**(S2) Interactive figures** I liked using the interactive figures provided in the paper. I especially liked Figure 2, which interactively demonstrates the potential problems in scatterplots. It really helped me gain an initial understanding of

the problem domain.

#### **Paper Major Weakness**

**(W1) Question on Novelty** The major weakness of the paper is that the proposed visualization system, which is Gatherplot, seems to have negligible novelty compared to previous works like Atom [2] or Kinetica [1]. Especially, except for GatherLens, the Gatherplot system seems to be almost identical to Atom Grammar. The paper should more precisely and explicitly describe the contribution of this paper, which differentiates it from previous papers.

**(W2) The Background section is too wordy and lacks connection to the proposed work** In the background section, there exists a list of methods that alleviate overplotting problems (e.g., appearance-based method, distortion-based method). However, the section is overwhelming (too many works listed), and the section also lacks connection to the proposed work. The paper should include how each type of method is related to the design of Gatherplot, and discuss the pros and cons of each type. This is especially important as the only competitor used in this paper is jittering.

**(W3) Section 3 (Gather transformation) is too complex and hard to understand** The description of the gather transformation, incorporating mathematical formulas, is hard to understand and seems to be unnecessarily complex. The mathematical notations are not reused, and the formulas are not referenced in the following sections. It would be better to explain the process in natural language, along with descriptive figures.

**(W4) The experiment needs more competitors** Though rigorously designed, the experiment needs to add more competitors. This limitation is also connected to W2. As the background section contains a rich set of previous works alleviating the overplotting, they should be included as competitors. Some of them are even easy to implement (e.g., transparency), so they can be easily added to the experiment.

**(W5) The experiment results lack discussions** The experiment results section should be more elaborate in terms of adding more interpretations and discussions. For example, when reporting the interaction effect, there needs to be following discussions on the effect and reasoning of the interaction effect.

**(W6) GatherLens needs evaluation** The GatherLens seems to be an interesting contribution, which adds substantial novelty to the paper (so that the paper can overcome W1). However, GatherLens itself needs a short evaluation that can confirm its usability. Maybe a simple usage scenario can help?

#### **Paper Minor Weaknesses**

**(M1)** In the Background – Data-aware Methods section, the paper claims that “Gatherplot requires no such balancing”, which differentiates it from Keim et al. [3]. Is this truly an advantage? It seems that there exists a natural tradeoff between

distortion and overlapping. Is it even possible to remove such balancing?

**(M2)** To facilitate the understanding of potential readers, it would be nice to illustrate how previous works (Background section) alleviate overlapping issues using figures.

**(M3)** Interval tick Marks → The paper is saying that the bracket is used because “it uses minimal ink and creates less density with adjacent ticks”. However, it seems not to be sufficient justification. Moreover, (f) seems to use more ink compared to (a), (d), and (e). Actually, I think the description of tick marks and its justification is better to be removed from the paper, as it incorporates minor design choices.

**(M4)** Personal interest: it would be better for Gatherplot to have the functionality to split groups into subgroups, enabling details-on-demand exploration of hierarchical data.

#### References

[1] Rzeszotarski, Jeffrey M., and Aniket Kittur. "Kinetica: Naturalistic multi-touch data visualization." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014.

[2] Park, Deokgun, et al. "Atom: A grammar for unit visualizations." IEEE transactions on visualization and computer graphics 24.12 (2017): 3032-3043.

[3] Keim, Daniel A., et al. "Generalized scatter plots." Information Visualization 9.4 (2010): 301-311.

---

### Openness/Transparency

The source code of the system is provided. However, the repository (<https://github.com/intuinno/gatherplot>) lacks detailed documentation. Still, the repository contains the way to run the project. I also appreciate that the online demo is provided. Overall, the openness and transparency of this system are not perfect, but they are still at a sufficient level to be published and announced.

---

### Submission categories

- Registered Report
- Replication Study
- Empirical Research - Quantitative
- Empirical Research - Qualitative
- Systems or design research
- Commentary
- Systematic Literature Review

---

### Suggested outcome

Major revisions: this paper requires substantial improvements that I will need to re-review to decide whether or not to endorse it.

---

## Requested changes

It would be great to see the next version in which all major and minor weaknesses are revised. At least the limitations listed in major weaknesses should be revised for publication.

---

## ORCID

*No response*

@codementum on  
Dec 20, 2023 20:33:

[referenced from [#\[DECISION\] Gatherplots 20-Nov-2023](#)]

@nickelm on  
Mar 24, 2024 20:49:

Verdict This paper presents an interesting visualization technique and the technique is rigorously validated through a large-scale crowdsourced user study. However, the paper currently holds several limitations (see major and minor weaknesses), and there exist questions on the novelty of the proposed system. Therefore, the paper currently needs substantial revision before publishing. Thus, the final decision is Major revision.

Thanks for the careful feedback. I hope that this new revision will address all the weaknesses.

Paper Strength (S1) Well-designed experiment The paper includes rigorous experiments using the Amazon Mechanical Turk platform. Due to its crowdsourced nature, the study results seem to be more credible. Despite there being some missing points in the study design (Check Weakness), the study is overall well-designed and executed.

Thank you.

(S2) Interactive figures I liked using the interactive figures provided in the paper. I especially liked Figure 2, which interactively demonstrates the potential problems in scatterplots. It really helped me gain an initial understanding of the problem domain.

Thank you. We are excited about this work as an interactive article.

Paper Major Weakness (W1) Question on Novelty The major weakness of the paper is that the proposed visualization system, which is Gatherplot, seems to have negligible novelty compared to previous works like Atom [2] or Kinetica [1]. Especially, except for GatherLens, the Gatherplot system seems to be almost identical to Atom Grammar. The paper should more precisely and explicitly describe the contribution of this paper, which differentiates it from previous papers.

All fair points. Obviously, Atom is our own work, even though Gatherplots preceded it and led to the formulation of the Atom grammar. However, we think there is a clear innovation here as the present paper focuses in particular on a specific research problem, presents the visual design, and evaluates the new technique compared to jittering scatterplots. We have added such wording to the paper.

**Revision:** We have clarified the distinction in a direct comparison at the end of the subsection titled "Data-aware Methods".

The point about Kinetica is well-taken. I was very aware of Kinetica, but always thought of it primarily as a forerunner for touch-based visualization as well as physical interaction for data visualization. Even if the focus of Kinetica is more on the natural data exploration using touch, there are clear parallels to gatherplots.

**Revision:** We have added a citation to Rzeszotarski and Kittur (2014) and compared it to gatherplots at the end of the Background section.

(W2) The Background section is too wordy and lacks connection to the proposed work. In the background section, there exists a list of methods that alleviate overplotting problems (e.g., appearance-based method, distortion-based method). However, the section is overwhelming (too many works listed), and the section also lacks connection to the proposed work. The paper should include how each type of method is related to the design of Gatherplot, and discuss the pros and cons of each type. This is especially important as the only competitor used in this paper is jittering.

I hear this feedback loud and clear. In response, I have removed the initial subsection titled "Characterizing Overplotting" on the basis that it was too basic and generally unnecessary.

**Revision:** We have reduced the Background section by eliminating the "Characterizing Overplotting" subsection.

(W3) Section 3 (Gather transformation) is too complex and hard to understand. The description of the gather transformation, incorporating mathematical formulas, is hard to understand and seems to be unnecessarily complex. The mathematical notations are not reused, and the formulas are not referenced in the following sections. It would be better to explain the process in natural language, along with descriptive figures.

This is very true. Other reviewers have noted the same.

**Revision:** We have eliminated unnecessary formalism and replaced it with natural language. In response, the Gather section is much shorter and easier to read.

(W4) The experiment needs more competitors. Though rigorously designed, the experiment needs to add more competitors. This limitation is also connected to

W2. As the background section contains a rich set of previous works alleviating the overplotting, they should be included as competitors. Some of them are even easy to implement (e.g., transparency), so they can be easily added to the experiment.

We are unable to add to the experiment, as the original students have long since graduated and left. While we agree that additional techniques would have been useful, we believe that the existence of the jittering condition is a fair baseline comparison.

(W5) The experiment results lack discussions The experiment results section should be more elaborate in terms of adding more interpretations and discussions. For example, when reporting the interaction effect, there needs to be following discussions on the effect and reasoning of the interaction effect.

Thanks for this feedback, it is very valid.

**Revision:** We have added a Discussion section and populated it with these types of findings.

(W6) GatherLens needs evaluation The GatherLens seems to be an interesting contribution, which adds substantial novelty to the paper (so that the paper can overcome W1). However, GatherLens itself needs a short evaluation that can confirm its usability. Maybe a simple usage scenario can help?

We appreciate the point. We were also quite enamored by the GatherLens example. However, based on this and other reviewer feedback, we have opted to remove it from the paper entirely.

**Revision:** The GatherLens section has been removed and is no longer claimed as a contribution.

Paper Minor Weaknesses (M1) In the Background – Data-aware Methods section, the paper claims that “Gatherplot requires no such balancing”, which differentiates it from Keim et al. [3]. Is this truly an advantage? It seems that there exists a natural tradeoff between distortion and overlapping. Is it even possible to remove such balancing?

**Revision:** The reviewer is correct. We have reformulated this passage.

(M2) To facilitate the understanding of potential readers, it would be nice to illustrate how previous works (Background section) alleviate overlapping issues using figures.

While this is a good idea and we entertained it, we found no good presentation for such a diversity of techniques. Thus, we opted not to do this for now.

(M3) Interval tick Marks → The paper is saying that the bracket is used because “it uses minimal ink and

creates less density with adjacent ticks". However, it seems not to be sufficient justification. Moreover, (f) seems to use more ink compared to (a), (d), and (e). Actually, I think the description of tick marks and its justification is better to be removed from the paper, as it incorporates minor design choices.

The point about ink is true; we have revised it to merely refer to "tick mark density". However, we have opted to keep the tick mark description in the paper since it adds to practical design aspects of the gatherplot technique.

**Revision:** We have revised the discussion about tick mark ink.

(M4) Personal interest: it would be better for Gatherplot to have the functionality to split groups into subgroups, enabling details-on-demand exploration of hierarchical data.

This is good feedback, but it is currently out of scope for our gatherplots implementation. We would point to the Dataopsy technique (Hoque & Elmqvist 2024; now cited), that supports some of this functionality.

Openness/Transparency The source code of the system is provided. However, the repository (<https://github.com/intuinno/gatherplot>) lacks detailed documentation. Still, the repository contains the way to run the project. I also appreciate that the online demo is provided. Overall, the openness and transparency of this system are not perfect, but they are still at a sufficient level to be published and announced.

We appreciate this feedback. Unfortunately, the original gatherplot implementation has been abandoned by the student and cannot be updated.

Major revisions: this paper requires substantial improvements that I will need to re-review to decide whether or not to endorse it.

Requested changes It would be great to see the next version in which all major and minor weaknesses are revised. At least the limitations listed in major weaknesses should be revised for publication.

Excellent, thank you. We believe we have addressed all of the major weaknesses and most of the minor ones. We appreciate the reviewer's careful feedback.

[@jov-anonymous-reviewer-AAAA-proxy](#) on May 14, 2024 08:14:

*Due to unknown reasons, I am currently not able to access the original account that I used to write the initial review. I thus created an alternative account (will report to editors)*

## Round 2 Decision

**Minor revisions:** this paper requires some smaller changes, after which I am confident I would be able to endorse it.

## Verdict

I would like to thank the authors for carefully addressing all the comments that I raised. I found that most of the serious concerns that existed in the initial document have been appropriately revised. I especially appreciate the authors for revising Section 3 (Gather transformation), which made the section much easier to understand and follow. I think that the paper will be ready for publication if only a few of the remaining weaknesses are addressed (Please see the weaknesses below; I tried my best to only tackle the weaknesses related to the ones that I previously raised). These weaknesses seem to be minor. Therefore, my decision is **minor revision**.

---

## Weaknesses

(W1) Still – why only “jitter”? I understand that it is hard to add more competitors to the experiment. I would not like to force authors to add competitors. Still, the reason why “only adding jittering as a baseline” is a valid experimental setting should be explained in detail. Without an explanation, potential readers may raise questions about the validity of the experimental setting.

(W2) Tick mark description – why not alternatives? The reason why another alternative design of tick marks is not selected should be explained in detail. I think the current description is still not convincing enough to justify the selection of the final design. This part needs more description. Also, it will be great to add how the alternative designs are selected.

(W3) Relation with Atom Thank you for adding Kinetica as related work and positioning the difference. Still, the relation with “Atom” grammar should be described in more detail. If the proposed system *led to the formulation of Atom grammar*, *in my humble opinion, such fact should also be noted. This is because the paper will anywhere be published after\** Atom paper. Therefore, readers may incorrectly think that Gatherplots is a future work of Atom (due to their chronological relation). To avoid such misinterpretation, the relationship with Atom should be detailed.

[@nickelm](#) on  
Jun 09, 2024 10:17:

---

## Weaknesses

(W1) Still – why only “jitter”? I understand that it is hard to add more competitors to the experiment. I would not like to force authors to add competitors. Still, the reason why “only adding jittering as a baseline” is a valid experimental setting should be explained in detail. Without an explanation, potential readers may raise questions about the validity of the experimental setting.

**Revision:** Thanks for this feedback. We have expanded the motivation in the "Experimental Design" subsection to carefully discuss all major classes of techniques and add rationale for the



choice.

(W2) Tick mark description – why not alternatives?

The reason why another alternative design of tick marks is not selected should be explained in detail. I think the current description is still not convincing enough to justify the selection of the final design. This part needs more description. Also, it will be great to add how the alternative designs are selected.

**Revision:** We have revised this section as well as the associated figure caption. The new treatment includes no unmotivated claims on visual clutter, and also discusses the origins of the given alternatives.

(W3) Relation with Atom Thank you for adding

Kinetica as related work and positioning the difference. Still, the relation with “Atom” grammar should be described in more detail. If the proposed system \*led to the formulation of Atom grammar”, in my humble opinion, such fact should also be noted. This is because the paper will anywhere be published **after** Atom paper. Therefore, readers may incorrectly think that Gatherplots is a future work of Atom (due to their chronological relation). To avoid such misinterpretation, the relationship with Atom should be detailed.

That the Gatherplots project led to the development of the Atom grammar is, in fact, already described in the paper (in a footnote). You can find the footnote in the last paragraph of the "Data-aware Methods" subsection. We hope that this note is sufficient for the reviewer.

@codementum on  
Jun 14, 2024 00:28:

[referenced from [#\[DECISION\] Gatherplots 13-May-2024](#)]

---