We plan to make the following changes:

1. (done) Add an explicit discussion on false positive and false negative results of both BayesDroid and TaintDroid. Emphasize that the accuracy hypothesis has been proven, as indicated by the accuracy field of Table 1 (Reviewers 1, 2, 3 and 4).
2. (done) Also, report on F-measure (a harmonized measure between precision and recall) for both BayesDroid and TaintDroid, which will further demonstrate that the accuracy hypothesis was proven (Reviewer 1).
3. (done) calculate the statistical significance and report on it
4. (done) Clarify the second hypothesis: comparison between dynamic taint tracking (T-BD) and an approach that compare sink and source values directly (H-BD). Explicitly calculate the accuracy of these approaches and explain the conclusion that H-BD was competitive with, if not better than, T-BD on the top-popular real-world applications we experimented with.
5. Extend the discussion on differences between the T-BD and H-BD approaches. Also, rephrase and elaborate on “more complete” findings and instabilities related to T-BD on p.11. (Reviewers 4 and 5).
6. (done) Refine Tables 1, 2 and 3 to present the current content in a more concise manner and include the accuracy calculations mentioned above. In addition, add more discussion of the tables to clarify the data they present (Reviewer 1).
7. Straighten the discussion on the threat model and the intended use of the technology. Specifically, the technology is not intended to defend against attackers who are aware of its use and thus have simple ways to circumvent it. It is rather intended to detect “accidental” releases of sensitive information by benign applications. (Reviewer 3).
8. Provide more explanations on the distance model. In particular, how it supports:
   * string transformations and encodings (Reviewer 2)
   * non-text data like bluetooth/microphone/camera (Reviewer 3)
   * multiple information units, e.g. for location (Reviewer 5)
9. Unify and clarify Sections 2.0 and 2.1 (Reviewer 3).
10. Add examples of conditional probabilities computed by our tool, and also list the full set of features (Reviewer 5).
11. Use |u| > |v| notations in Algorithm 1 instead of Min and Max (Reviewer 5).
12. Clarify that the “values reachable from sink arguments” statement on p.10 indeed talks about information flow (Reviewer 5).
13. Rephrase/drop the "Motivation" paragraph on p.10 (Reviewer 5)
14. Add axis labels (with units) and legend to Figure 5 (Reviewers 1 and 3).
15. (done) Remove repetitions about the deployment of a system in a commercial product (Reviewers 2 and 3).
16. (done) Fix typos:
    * "we used to for our estimates" (page 6);
    * "smoothening" (page 8) - should be "smoothing";
    * also on page 8, "to avoid from zero" - should be "to avoid zero".
    * \color{purple} in Fig.4
17. Add the following to the related work section
    * Zozzle, USENIX Security 2012 (Reviewer 2)
    * R. Sekar, "An Efficient Black-box Technique for Defeating Web Application Attacks", NDSS 2009 (Reviewer 5)
    * A relationship with reflected XSS filters in web browsers (Reviewer 5)

We do not plan to:

1. Add a comparison of the runtime performance between BayesDroid and TaintDroid (Reviewer 1).

Explanation: Such a comparison is difficult to perform in a reliable manner due to the interactive (UI-driven) nature of mobile applications.

1. Modify the Bayesian approach to update the probability inference when collecting more evidence (Reviewer 2).

Explanation: This is an interesting suggestion, which we are thankful for. We view it as an extension of our algorithm, which goes outside its current scope and evaluation, and we will address it in future research.

1. Remove definition of Levenshtein distance and Algorithm 1 (Reviewer 3).

Explanation: The definition requires little space, and forms an essential part of our approach and algorithm.