================================================================  
  
  
        --========  Review Reports  ========--  
  
The review report from reviewer #1:  
  
\*1: Is the paper relevant to ICDM?  
  [\_] No  
  [X] Yes  
  
\*2: How innovative is the paper?  
  [\_] 6 (Very innovative)  
  [\_] 3 (Innovative)  
  [X] -2 (Marginally)  
  [\_] -4 (Not very much)  
  [\_] -6 (Not at all)  
  
\*3: How would you rate the technical quality of the paper?  
  [\_] 6 (Very high)  
  [\_] 3 (High)  
  [X] -2 (Marginal)  
  [\_] -4 (Low)  
  [\_] -6 (Very low)  
  
\*4: How is the presentation?  
  [\_] 6 (Excellent)  
  [\_] 3 (Good)  
  [X] -2 (Marginal)  
  [\_] -4 (Below average)  
  [\_] -6 (Poor)  
  
\*5: Is the paper of interest to ICDM users and practitioners?  
  [\_] 3 (Yes)  
  [\_] 2 (May be)  
  [X] 1 (No)  
  [\_] 0 (Not applicable)  
  
\*6: What is your confidence in your review of this paper?  
  [X] 2 (High)  
  [\_] 1 (Medium)  
  [\_] 0 (Low)  
  
\*7: Overall recommendation  
  [\_] 6: must accept (in top 25% of ICDM accepted papers)  
  [\_] 3: should accept (in top 80% of ICDM accepted papers)  
  [\_] -2: marginal (in bottom 20% of ICDM accepted papers)  
  [X] -4: should reject (below acceptance bar)  
  [\_] -6: must reject (unacceptable: too weak, incomplete, or wrong)  
  
\*8: Summary of the paper's main contribution and impact  
  The paper is an extension of the work published ACM TKDD in 2009. The clarity of extension to the work, and contributions are below par.  
  
\*9: Justification of your recommendation  
  The authors state - we yield summaries that are accurate and concise and correct. I tried to find definitions, and results on this in the paper, it is not there.  
  
The notion of local and its evolution to global patterns is there, and the aspect of optimality is there. But whether the optimal solution is correct and accurate is not clear.  
  
\*10: Three strong points of this paper (please number each point)  
  1. The problem is a generalization of existing work  
2. MDL oriented optimization.  
3. Some results on real data sets.  
  
\*11: Three weak points of this paper (please number each point)  
  1. Results are misleading as we do not know what is accurate and correct.  
2. MDL optimization motivation is not clear as to whether it formulates the problem correctly.  
3. The grouping of events done in this manner may lose out important details. Such as finite state modeling of even generators.  
  
\*12: Is this submission among the best 10% of submissions that you reviewed for ICDM'18?  
  [X] No  
  [\_] Yes  
  
\*13: Would you be able to replicate the results based on the information given in the paper?  
  [X] No  
  [\_] Yes  
  
\*14: Are the data and implementations publicly available for possible replication?  
  [X] No  
  [\_] Yes  
  
\*15: If the paper is accepted, which format would you suggest?  
  [\_] Regular Paper  
  [X] Short Paper  
  
\*16: Detailed comments for the authors  
  Segmentation problem is age old, and as applied to event streams is not novel. Figure 2 example, is it pre-decided that those intervals have those events. Otherwise, I thought the problem was to find the event sequences in time lines. Therefore, it is not clear whether sudden increase in numbers of the matrix after 8, and then decrease in numbers for e1 after 17, played any part in the solution got.   
  
  
The above example seems to be trivial. It has to be seen how the algorithm works, when e1 occurs at every time unit i, e2 occurs at every odd time unit, and e3, occurs at every prime number unit. In this, case what will the algorithm output. What interval segmentations will it perform.  
  
So, the issue with the paper, is rationale for the approach, and substantial lack of effort in convincing that the approach works. Also, not comparing with accurate and correct solution, makes the paper weaker.  
  
Is it a lossy compression you get?  
  
Further there is work by Sharma Chakravarthy on event composition, and detection, check Snoop work. Is that not a better way to represent event sequences.

========================================================  
The review report from reviewer #2:  
  
\*1: Is the paper relevant to ICDM?  
  [\_] No  
  [X] Yes  
  
\*2: How innovative is the paper?  
  [\_] 6 (Very innovative)  
  [X] 3 (Innovative)  
  [\_] -2 (Marginally)  
  [\_] -4 (Not very much)  
  [\_] -6 (Not at all)  
  
\*3: How would you rate the technical quality of the paper?  
  [\_] 6 (Very high)  
  [X] 3 (High)  
  [\_] -2 (Marginal)  
  [\_] -4 (Low)  
  [\_] -6 (Very low)  
  
\*4: How is the presentation?  
  [\_] 6 (Excellent)  
  [X] 3 (Good)  
  [\_] -2 (Marginal)  
  [\_] -4 (Below average)  
  [\_] -6 (Poor)  
  
\*5: Is the paper of interest to ICDM users and practitioners?  
  [\_] 3 (Yes)  
  [X] 2 (May be)  
  [\_] 1 (No)  
  [\_] 0 (Not applicable)  
  
\*6: What is your confidence in your review of this paper?  
  [X] 2 (High)  
  [\_] 1 (Medium)  
  [\_] 0 (Low)  
  
\*7: Overall recommendation  
  [\_] 6: must accept (in top 25% of ICDM accepted papers)  
  [X] 3: should accept (in top 80% of ICDM accepted papers)  
  [\_] -2: marginal (in bottom 20% of ICDM accepted papers)  
  [\_] -4: should reject (below acceptance bar)  
  [\_] -6: must reject (unacceptable: too weak, incomplete, or wrong)  
  
\*8: Summary of the paper's main contribution and impact  
  The authors presented a formulation and an optimal strategy to summarize binary valued time series. The method yielded concise and accurate summaries and provided a global data model. Additionally, this method provided a local pattern model and was parameter-free.  
  
\*9: Justification of your recommendation  
  The authors present a technical paper with relevant topic and proper research methodology. I consider that it is a very interesting paper and matches well to the conference session.  
  
\*10: Three strong points of this paper (please number each point)  
  1. The paper includes the up-to-date references.  
2. The technical approach and analysis is very rigid.  
3. The paper is well-written and its structure is well organized.  
  
\*11: Three weak points of this paper (please number each point)  
  1. The labels of the graph are not clear.  
2. The performance of the real-world dataset should be compared with the baseline model.   
3. There is something confusing in the style and grammar.  
  
\*12: Is this submission among the best 10% of submissions that you reviewed for ICDM'18?  
  [X] No  
  [\_] Yes  
  
\*13: Would you be able to replicate the results based on the information given in the paper?  
  [\_] No  
  [X] Yes  
  
\*14: Are the data and implementations publicly available for possible replication?  
  [X] No  
  [\_] Yes  
  
\*15: If the paper is accepted, which format would you suggest?  
  [\_] Regular Paper  
  [X] Short Paper  
  
\*16: Detailed comments for the authors  
  This paper is acceptable and interesting and fits within the scope of the conference.   
However, some modifications should be made to improve the paper, as follows:  
  
- The labels of the graphs are confusing. It is suggested to re-demonstrate the labels of each graph.   
  
- The performance of the real-world dataset should be compared with some baseline model. How about try some other algorithms and see the difference?   
  
- In the beginning of the experiment evaluation, the authors should change the grammar of the sentences. For example, "Will the model find useful patterns.." instead of "The model finds useful patterns from the real world dataset?"

========================================================  
The review report from reviewer #3:  
  
\*1: Is the paper relevant to ICDM?  
  [\_] No  
  [X] Yes  
  
\*2: How innovative is the paper?  
  [\_] 6 (Very innovative)  
  [\_] 3 (Innovative)  
  [X] -2 (Marginally)  
  [\_] -4 (Not very much)  
  [\_] -6 (Not at all)  
  
\*3: How would you rate the technical quality of the paper?  
  [\_] 6 (Very high)  
  [\_] 3 (High)  
  [X] -2 (Marginal)  
  [\_] -4 (Low)  
  [\_] -6 (Very low)  
  
\*4: How is the presentation?  
  [\_] 6 (Excellent)  
  [\_] 3 (Good)  
  [\_] -2 (Marginal)  
  [X] -4 (Below average)  
  [\_] -6 (Poor)  
  
\*5: Is the paper of interest to ICDM users and practitioners?  
  [X] 3 (Yes)  
  [\_] 2 (May be)  
  [\_] 1 (No)  
  [\_] 0 (Not applicable)  
  
\*6: What is your confidence in your review of this paper?  
  [X] 2 (High)  
  [\_] 1 (Medium)  
  [\_] 0 (Low)  
  
\*7: Overall recommendation  
  [\_] 6: must accept (in top 25% of ICDM accepted papers)  
  [\_] 3: should accept (in top 80% of ICDM accepted papers)  
  [\_] -2: marginal (in bottom 20% of ICDM accepted papers)  
  [X] -4: should reject (below acceptance bar)  
  [\_] -6: must reject (unacceptable: too weak, incomplete, or wrong)  
  
\*8: Summary of the paper's main contribution and impact  
  The paper presents an approach using Minimum Description Length and dynamic programming to summarize/compress a set of different time series (frequencies of different event types) by partitioning the timeline into segments, grouping the subsequences in each partition and representing them using a simple model (average number of events). In addition to compressing the data, the method may also reveal patterns in the data via the local clustering of sequences in the segments.  
  
\*9: Justification of your recommendation  
  The method seems to be a rather straightforward generalization of reference [1] in the paper (using Poisson distributions to model event frequencies), but the way it is presented rather obscures than clarifies this fact. The experimental evidence for the usefulness of the approach is also not very strong.  
  
\*10: Three strong points of this paper (please number each point)  
  1) Based on mathematical principles.   
2) Good use of illustrations to help explain the concepts.  
3) Uses a real data set in the evaluation.  
  
\*11: Three weak points of this paper (please number each point)  
  1) Notation messed up in several places.  
2) Experimental evaluation weak.  
3) Not well written.  
  
\*12: Is this submission among the best 10% of submissions that you reviewed for ICDM'18?  
  [X] No  
  [\_] Yes  
  
\*13: Would you be able to replicate the results based on the information given in the paper?  
  [X] No  
  [\_] Yes  
  
\*14: Are the data and implementations publicly available for possible replication?  
  [X] No  
  [\_] Yes  
  
\*15: If the paper is accepted, which format would you suggest?  
  [\_] Regular Paper  
  [X] Short Paper  
  
\*16: Detailed comments for the authors  
  The notation is inconsistent and even confusing in some places: In equation (1), the code length, which is a number, is identified with a model, and M\_i is equated with an expression that contains M\_i (M\_i = argmin\_{M'\_i} (L(M\_i) + …). Why have L and LD to denote the same thing in equation (3) and (5)? (There is also one occurrence of \lambda\_{daily}, which should be \lambda\_d.  
  
The experimental evaluation is weak: 1) It does not describe in enough detail how the synthetic data is generated (e.g. what exactly did you do to "create 30 correlation patterns with 30 defined segments"?). I am assuming that the data is generated corresponding to the exact model assumptions that you make in your method, so that part of the evaluation essentially shows only that you have implemented your proposal correctly - or could there have been any surprises? The evaluation on the real data also does not provide much insight, and it is not clear whether the simple results/observation could not have been made with much simpler approaches - that also points to the lack of any comparison with other approaches or baseline methods.   
The language is quite rough in several places and should be improved.   
  
Section III, while interesting on its own, is not really relevant for presenting the content of the paper. Using MDL is a choice, not a necessity! It has its advantages and drawbacks compared to other approaches to compress/summarize time series using models (e.g., driven by an error bound). Those could be discussed instead (and evaluated against in the experimental section!).  
========================================================  
:  
  
Meta Review:  
  
Two reviewers observed that the paper is a marginal extension of an existing paper [1] and that it is not clearly stated what the novel contributions are. Adding this to the weak evaluation and writing, this unfortunately makes it impossible to consider the paper for publication at ICDM.