在线评论中不文明行为的时间模式研究报告

背景信息 (Background Information)

数字媒介互动（如在线评论区）中的不文明行为是一个严重的社会问题，对个人福祉和社会凝聚力构成威胁。它可能导致心理健康受损、社会两极分化加剧，并对在线民主讨论产生负面影响。据估计，不文明帖子的比例在2%到20%之间。在新闻网站的评论区，不文明行为不仅影响新闻机构的可信度，还可能导致记者因担心失去消息来源而感到恐惧，并加剧错误信息的传播。问题的严重性甚至促使一些新闻机构完全关闭了评论功能。

尽管已有大量关于不文明行为的研究，但很少有研究探讨其发生的速度与可能性之间的直接关系。本研究旨在填补这一空白，探讨评论发布的速度是否会影响其文明程度。研究假设，快速发布的评论更可能是不文明的，并且不文明行为会通过“传染效应”进一步扩散。

理论框架与研究模型 (Theoretical Framework/Research Model)

本研究的理论框架主要建立在两个核心概念之上：对“不文明行为”的界定，以及认知心理学中的双系统理论（Dual-Process Theory）。

不文明行为的界定

研究首先对“不文明行为”（incivility）进行了界定。不文明行为被视为一种违反社会礼仪规范的“反规范”交流。文献中对不文明行为的定义并不统一，但通常可以分为两大类：

1. **严重违规行为：** 这类行为会威胁民主原则，例如针对种族、性别、性取向等受保护特征的攻击，或直接的人身威胁。Papacharissi (2004) 将此类行为定义为“不文明”（incivility）。
2. **轻度违规行为：** 这类行为虽然粗鲁无礼，但不会动摇民主价值观，例如侮辱、尖锐的批评或轻蔑的言论。Papacharissi (2004) 将其称为“不礼貌”（impoliteness）。

本研究采纳了Papacharissi (2004) 的区分，将研究焦点放在前者，即具有严重后果的“不文明行为”上。这一选择与研究的数据来源——《卫报》的社区标准相契合，因为《卫报》的版主在删除评论时，主要针对的是仇恨言论和威胁性行为，而非单纯的粗鲁言辞。这种界定为研究提供了一个清晰且可操作的测量标准。

双系统理论与在线评论

研究的核心理论模型借鉴了诺贝尔奖得主丹尼尔·卡尼曼（Daniel Kahneman, 2011）提出的双系统理论，该理论将人类的认知过程分为两种模式：

* **系统1（快速思考）：** 这是一种直觉、自动且毫不费力的思维过程。它通常与情绪化、本能的反应相关联。例如，在面对冒犯性言论时，人们的第一反应通常由系统1主导。
* **系统2（慢速思考）：** 这是一种需要投入精力、更具逻辑性和审慎性的思维过程。它负责复杂的计算、逻辑推理以及对自身行为的持续监控。卡尼曼指出，保持礼貌，尤其是在愤怒时，是系统2发挥作用的典型例子，因为它需要刻意控制自己的冲动。

本研究将该理论应用于在线评论行为，提出**不文明评论是“快速思考”（系统1）的产物**。当用户受到激发（例如，阅读了一篇有争议的文章或一条挑衅性的评论）时，其本能的、情绪化的反应（系统1）可能会立即转化为不文明的言论。相比之下，**文明的、有建设性的讨论则需要“慢速思考”（系统2）的参与**，用户需要花时间深思熟虑，构建更有逻辑的论点，并抑制不当言论的冲动。

基于此模型，研究提出了两个核心假设：

* **假设1：** 不文明的在线评论比文明的评论发布得更快，因为它们是系统1主导的即时反应。
* **假设2：** 不文明的交流会引发“传染效应”，即一条不文明评论会激发其他用户更快速、更情绪化的反应，从而导致更多不文明评论的出现。

研究方法 (Research Methodology)

本研究旨在探讨在线评论中不文明行为的时间模式，并验证其与评论速度及传染效应之间的关系。为此，研究采用了大规模数据集分析，并结合多种统计方法对数据进行处理和解释。

数据收集

本研究的数据来源于《卫报》（The Guardian）网站的在线新闻评论语料库，具体聚焦于其“观点”(Opinion)版块。该数据集涵盖了自2006年3月至2024年3月期间发布的约3800万条读者评论。研究将《卫报》版主删除的评论作为“不文明行为”的代理（proxy），共计约100万条。选择《卫报》作为数据来源，是基于其庞大的用户基础、活跃的评论社区以及明确的社区管理标准，这为研究提供了丰富的真实世界数据和对不文明行为的明确界定。数据中包含详细的时间戳元数据，这对于分析评论速度和时间模式至关重要。

关键变量的操作化与测量

1. **不文明行为（Incivility）**：本研究将不文明行为操作化定义为《卫报》版主删除的评论。这一操作化定义与Papacharissi (2004)的理论框架相符，该框架将具有严重社会影响（如仇恨言论、威胁行为）的交流视为不文明行为，并将其与相对轻微的“不礼貌”行为区分开来。通过这种方式，研究能够利用《卫报》现有的、基于其社区标准的分类，获取大量已被识别为不文明的评论样本。
2. **评论响应时间（Commenting Time）**：评论响应时间是本研究的核心时间变量，用于衡量评论发布的快慢。其计算方式分为两种：
   * **父评论响应时间**：对于直接回复文章的评论，其响应时间定义为评论发布时间与文章发布时间之间的时间间隔。
   * **子评论响应时间**：对于回复其他评论的评论（即子评论），其响应时间定义为子评论发布时间与所回复的父评论发布时间之间的时间间隔。

数据分析方法

鉴于评论响应时间数据通常不符合正态分布，本研究主要采用了非参数统计方法进行数据分析，以确保结果的稳健性和准确性。具体分析方法包括：

1. **Mann-Whitney U检验**：该检验用于比较两组独立样本（例如，被版主删除的不文明评论与未被删除的文明评论）中位数是否存在显著差异，以评估不文明评论是否确实发布得更快。
2. **Kruskal-Wallis检验**：当需要比较三个或更多独立样本组之间的差异时，研究采用了Kruskal-Wallis检验。若该检验结果显著，则会进一步使用Dunn事后检验（post hoc test）进行两两比较，以确定具体哪些组之间存在差异。
3. **Wilcoxon符号秩检验**：此检验用于比较成对样本的差异，例如在分析评论链中不文明行为的传播模式时，可能用于比较同一评论者在不同情境下的评论速度。
4. **Spearman等级相关系数**：为了量化不文明行为传染效应的强度，研究使用了Spearman等级相关系数来衡量时间间隔与后续评论被屏蔽概率之间的关系。这有助于理解不文明行为如何在评论序列中扩散。
5. **自助法（Bootstrapping）**：为增强统计推断的可靠性，研究还采用了自助法来计算中位数差异的置信区间，并通过随机分组置换来验证结果的稳健性。这对于处理大规模和复杂的数据集尤为重要。

研究假设

本研究基于上述理论框架和数据分析方法，提出了以下两个核心假设：

* **假设1**：不文明的在线评论比文明的评论发布得更快。反之，文明的评论在发布前有更长的时间间隔，这与“快速思考”和“慢速思考”的认知过程相对应。
* **假设2**：不文明的交流会增加后续出现更多不文明交流的可能性，即存在不文明行为的“传染效应”。

研究发现与局限性 (Findings and Limitations)

本研究主要探讨了在线评论中不文明行为的时间模式及其传染效应，并得出了以下主要发现：

主要发现 (Findings)

1. **不文明评论发布速度更快**：
   * **子评论**：被屏蔽的子评论（代表不文明评论）的发布速度显著快于可见的子评论。具体而言，被屏蔽的子评论的中位时间为47分钟，而可见子评论的中位时间为15分钟。这表明不文明的回复往往是即时且冲动的。
   * **父评论**：类似地，被屏蔽的父评论（直接回复文章的评论）的发布速度也快于可见的父评论。被屏蔽的父评论中位时间为176分钟，而可见父评论为58分钟。
   * **扁平评论**：对于扁平评论（flat comments），即不属于任何评论串的独立评论，被屏蔽的评论同样比可见评论发布得更快。
   * 这些发现支持了研究的第一个假设，即不文明行为与“快速思考”相关联，而文明交流则需要更长的“慢速思考”时间。
2. **不文明行为的传染效应**：
   * 研究发现，当一个父评论至少包含一个被屏蔽的子评论时，后续的子评论被屏蔽的可能性会显著增加。具体来说，在评论串中，如果一个父评论有一个被屏蔽的子评论，那么后续的第二个子评论被屏蔽的概率为13.66%，而第六个子评论被屏蔽的概率高达49.92%。
   * 这种效应在评论串中持续存在，且随着评论链的深入，不文明行为出现的可能性及其强度（以被屏蔽评论的数量衡量）均呈上升趋势。
   * 此外，不文明评论的出现还会缩短后续评论的响应时间，进一步证实了不文明行为的传染性，即一条不文明评论会引发更快速、更情绪化的反应，从而导致更多不文明评论的出现。这支持了研究的第二个假设。

研究局限性 (Limitations)

本研究虽然取得了重要发现，但也存在一些局限性：

1. **不文明行为的代理测量**：研究将《卫报》版主删除的评论作为不文明行为的代理。尽管这与《卫报》的社区标准相符，且提供了大规模的真实世界数据，但这种操作化定义可能无法捕捉所有形式的不文明行为。例如，一些不文明但未被删除的评论可能被遗漏，而一些被删除的评论可能并非严格意义上的“不文明”，而是违反了其他社区规定。
2. **因果关系推断**：尽管研究揭示了评论速度与不文明行为之间的强关联，以及不文明行为的传染效应，但由于数据是观察性的，难以完全排除其他潜在的混淆变量，因此在严格的因果关系推断上仍需谨慎。
3. **平台特异性**：本研究的数据来源于《卫报》这一特定新闻平台。尽管《卫报》是一个具有代表性的在线新闻媒体，但其用户群体、评论文化和审核政策可能与其他平台存在差异。因此，研究结果的普适性可能受到一定限制，未来研究可以在不同平台进行验证。
4. **时间粒度**：研究使用的时间戳数据可能存在一定的粒度限制，虽然足以揭示宏观的时间模式，但对于极其微秒级的评论行为差异可能无法完全捕捉。
5. **缺乏用户个体信息**：数据集主要关注评论内容和时间戳，缺乏用户的详细人口统计学信息或心理特征数据。这限制了研究深入探讨用户个体差异如何影响不文明行为及其时间模式的能力。

结论与未来研究方向

本研究的发现强调了在线评论中“快速思考”与不文明行为之间的紧密联系，以及不文明行为的传染性。这些结果对于在线平台设计具有重要启示，鼓励平台通过引入“审慎时间”（deliberative time）来促进更健康的讨论环境。未来的研究可以尝试结合更精细的用户行为数据、进行实验设计以验证因果关系，并在不同文化和平台背景下拓展研究，以更全面地理解在线不文明行为的动态。