

AirDiary : Bluetooth デバイス検出履歴を用いた半自動日記作成ツール

牛越達也*† 河野恭之†

本研究は、日常的に自身の周辺に存在する Bluetooth デバイスを検出して得られる Social context を解析することで短文形式の日記を半自動的に作成する UI を作成した。Bluetooth デバイスログより他者との出会い、共在の特徴を捉え、長期的に観測することで日常の中に潜む非日常的な出来事をユーザと協調的に検出する。また、検出され日常、非日常の特徴を学習することで一度検出した出来事は、自動に検出が可能である。そうして、検出された日常的、非日常的な出来事を基に、テンプレート方式で短文の日記を作成する。また、twitter と連携することで日記の公開、共有が可能である。

1. はじめに *

日記は、古来より日々の出来事の記録としてたしなまれてきた。その大きな目的は、自身が体験した出来事をありのままに記録する記憶の補助や、自身の“人生”を他者に公開することで文章を介して興味情報の交換を行うためである。記憶の補助は、自身が忘れたくないかけがえの無い記憶を当時の新鮮な感覚と共に文章化し、外部記憶に預けることで永遠に残そうとする人の回帰願望が具体化したものである。また、そうした日記を他者に公開することで、読者に自身の人生との追体験による“共感”や筆者が体験した出来事に対する“興味”を生み出す。近年、そうした日記の創作活動の場が大きな広がりを見せている。個人で開設する Blog や mixi, Facebook などの SNS の発達に伴い日記を web に記録にすることが容易になったことで、web で繋がりをもつ個人とのコミュニケーション手段として日記が公開されている[1][2]。こうした web がもたらした人間関係の変化や、手軽に記録と共有ができる環境によって日記作成の重要性は高まっている。しかし、日記作成は全ての人が容易にできるわけではない。多くの人が日記作成の手間を“面倒”と感じ、加えて体験を思い起こすことが苦手な人には不可能ともいえる過程がある。日記作成には、次のような過程が必要である。

I. 日々の中から日記に書くべき出来事を見つける

II. 記憶を再構成する

III. 文章を書く

I は、日々起きた出来事を思い返し、書きたいと思える出来事を見つけることである。書きたい出来事は、帰宅、仕事など頻繁に生じる出来事すなわち日常ではなく、日常には起きない出来事すなわち非日常から選ばれることが多い。II では、I で見つけた出来事の記憶を手がかりに詳細を思い出し、当時の感覚を呼び戻す。そのため、I と II の段階で日記の面白さと書くモチベーションが決まる。しかし、体験した思い出を覚えておくことが苦手な人にとって I, II は困難が伴う。体験した思い出を覚えておくことが苦手な人は、時間がたった出来事を思い出すことが困難であり、I で面白い出来事を見つけられないために II の過程に至ることができない。また、III に限っては多くの人が“面倒”と感じるだろう。そのため、日記を書きたい願望があっても多くの人が行動に移せないでいる。そこで、本研究では日記を書くことを“面倒”と感じる人たちのために、AirDiary と呼ぶ半自動的に日記作成支援するツールを提案する。AirDiary は、モバイル端末を通じて取得した Bluetooth デバイス検出履歴(BT ログ)を基に自身と他者の共在を解析し、日常的、非日常的な出来事をユーザに提示することで書きたい出来事の探索を支援する。また、日常的、非日常的な出来事をユーザに提示し、予め用意したテンプレートに選択した出来事を埋め込むことで短文形式の日記を作成する。また、twitter[3]と連携することで作成した日記を公開、共有する。

日記の作成支援を目的とした研究がある。坂本らは、個人の経験や興味を漫画で表現するコミックダイアリーを提案している[4]。コミックダイアリーは、博物館や学会など特定の出来事に専門的に機能するものであり、専用のガイド端末を利用して閲覧履歴や興味情報を蓄積し、知的処理によってストーリーを作成する。作成されたストーリーを基にテンプレートを組み合わせ、コミック形式の日記を作成するものである。閲覧履歴やユーザの行動履歴を用いることでストーリーを形成し、コミック特有の誇張表現によって興味をダイレクトに伝える手法としては優れているが、特定の出来事に対して専用のテンプレートを用意するなど日々の出来事を対象として日記の作成支援をするシステムではない。NTT レゾナント、NTT コミュニケーションズが提供している GPS 行動履歴を活用した日記作成サービスであるキセキがある[5]。キセキは、GPS 搭載携帯を利用することで訪れた場所の住所と時間を自動的に記録し、その記録に対してコメントを付けることで行動日記を作成する。GPS を用いることでショッピングなど場所に関連付く出来事の日記は作成可能であるが、場所移動が無い出来事に対する日記の作成は困難である。そこで、本研究は BT ログを基に自身と他者の織り成す周辺環境の変化をとらえ、場所移動に影響されない日常的、非日常的な出来事を検出することで日記の作成支援をする。

*† 関西学院大学大学院
Kwansei Gakuin University Graduate School

2. 本研究のアイデア

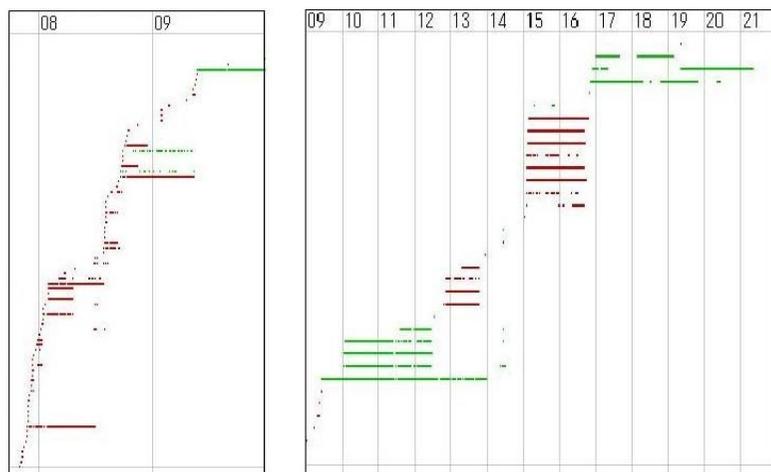


図1 BTログ (左: 電車 右: 大学内)

本研究のアイデアは BT ログを長期的に観測することで、周期的に生じる Bluetooth Device Address(BDA)の組み合わせを捉え、日常と非日常を検出することである。また、検出した出来事を基にテンプレート形式で日記を作成する。日記に記録する出来事は、日常的な出来事ではなく非日常的な出来事である。日常とは、大学の講義、仕事など一定の周期をもつ出来事である。逆に、非日常とは突然の来訪、休講など周期性がない出来事である。そのため、日常、非日常を検出するには、なんらかの特徴を基に周期的に生じる特徴を捉える必要がある。

本研究では日々の出来事を解析する情報源として Bluetooth(BT)に注目した。BT は、近距離無線通信の規格でありノート PC, PDA, 携帯電話など人に付随して動く機器の多くに搭載されている。そのため、BT 機器を認識することは自身と他者との“出会い”を検出している。また、BT 機器には 48 ビットで表現される BDA と呼ばれるユニークな ID が付与されているために識別が可能である。図 1 の左は電車、右は大学内で取得した BT ログを可視化したものである。上部の数字は時刻(24 時間表記)、縦軸は得られた BDA に検出順に付与した ID である。左は 8 時から 9 時の間に同じ電車かつ同じ車両内で、被験者が取得した BT ログである。車両内の人の入れ替わりや、同じ車両に乗っている他者との共在が同一の BDA の継続で示される。また、右は大学内での生活で取得された BT ログである。この BT ログが取得された日は、10 時から 12 時 30 分まで被験者を含め 6 人で会合をした後、6 人のうち 1 人と 12 時 50 分ごろから 13 時 50 分

ろまで食事に出かけている。そのため、ある BDA が 12 時 50 分ごろから 13 時 50 分ごろまで継続して検出されている。これらの特徴は、場所に影響されずに自身と他者との共在で日々の出来事を表している。

BT 機器を用いて行動を解析する研究がある。Nicolai らは、特定可能な人物の BT 機器の同時検出数に注目することで周辺環境の変化を検出している[6]。Kostakos らは、BT 機器が共存することを周辺環境として捉え、多くの人々が共存する場所の特定に用いている[7]。Donnie は、自身の周辺に存在する BT 機器の検出数を周辺環境とすることで、共存する人の粗密に着目している[8]。彼らは、自身の周辺に存在する特定もしくは不特定の BT 機器の検出数を周辺環境とすることで共存する BT 機器の疎密に着目し、行動履歴の推定をしている。しかし、ユーザの行動の変化を捉えるのには BT 機器の粗密は重要ではあるが、行動の推定や分類をするためには BDA の継続的な共在こそが重要である。そこで AirDiary は、継続して検出される BDA に注目し、一定の周期で記録されている BDA の組み合わせを、日常に関連するものとして日常の検出に用いる。非日常は、日常として判定されなかった時間帯に注目することで、ユーザと協調的に検出する。

3. AirDiary

AirDiary は、個人の利用に特化したアプリケーションであり、ユーザが常時取得した BT ログより日常的、非日常的な出来事をユーザと協調的に検出する。また、検出された日常に対してユーザにイベントタグの付与を求め、以後一度検出された日常を自動で検出し、イベントタグを付与する。更に、検出した出来事を基にテンプレート形式の日記を作成する。図 2 は、AirDiary の内部処理をフローで示したものである。日常の検出には、共起履歴と教師データを用いる。日常は、大学の講義、自宅など一定の周期で生じる出来事であり、BT ログからも周期的に継続的に生じる BDA の組み合わせとして観測される[9]。共起履歴は、長期間蓄積された BT ログを解析することで周期的に生じる継続的に検出される BDA の組み合わせを登録したものである。AirDiary は、日常、非日常を検出しユーザに対して正誤を求め、問いに対するユーザの修正情報を修正履歴に蓄積し学習に用いる。教師データは、修正履歴を学習することで、一度検出された出来事に関連する最適な BDA の組み合わせを登録したものである。非日常は、日常として判定されなかった時間帯にある可能性が高い、そのため非日常の可能性のある時間帯をユーザに提示し、非日常の正誤を求めることで非日常的な時間帯を検出する。

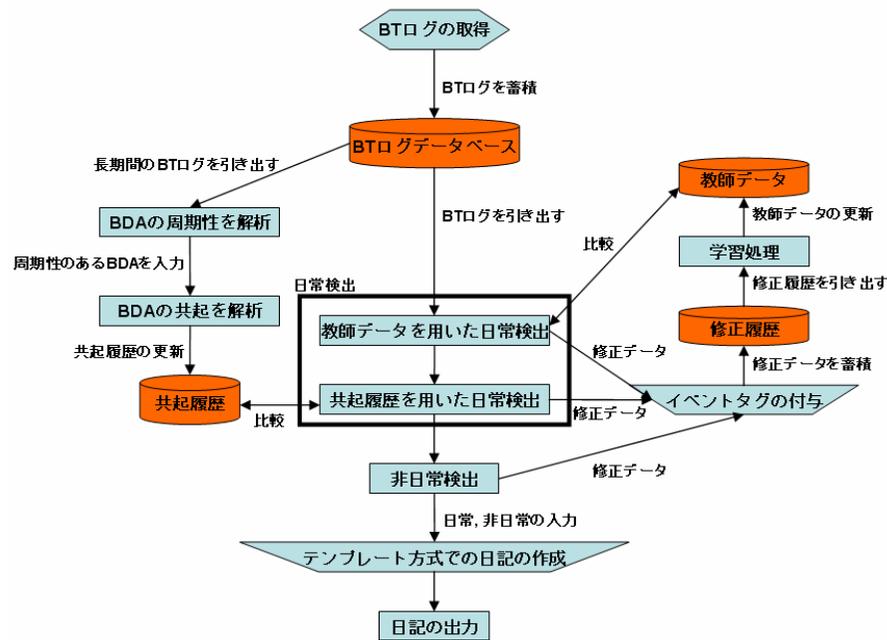


図 2. AirDiary の処理フロー

4. 共起履歴と教師データを用いた日常の検出

4.1 共起履歴の作成

BT ログから出来事を検出するためには、BDA の継続的な共起を検出することが重要である。また、日常的に生じる出来事に関連する BDA の組み合わせは周期的に繰り返し検出されると考えられる。そのため、長期にわたり周期的に検出される BDA に注目して共起しやすい BDA の組み合わせを導出する。解析に用いる BT ログには取得環境やプロトコルの特性により電波到達距離内にある BT 機器でも検出漏れが頻繁に生じるため、ガウシアンフィルタで補間したものを用いる。以下にその手順を示す。

Step1. 周期性をもつ BDA の抽出

継続的に検出される BDA の継続開始時刻に注目し、日が変わってから再び継続的

に検出されるまでの時間を基にヒストグラムを作成する。以後これを周期性ヒストグラムと呼ぶ。図 3 は、ある被験者の長期間の BT ログより周期性ヒストグラムを作成した例である。横軸は継続する BDA が生じてから日が変わって再び生じるまでの時間、縦軸は同じ周期に属する BDA の数を示している。24 時間、168 時間の周期に多くの BDA が検出されており、日常的に繰り返し検出される BDA が確認される。周期性ヒストグラムは、急な予定の変化や日常的に起きる行動の変化によって、常に一定の周期になるとは限らない。そうした周期のずれを許容するために、24 時間単位の周期を中心にその周囲のデータを寄せ上げる処理をする。寄せ上げた周期性ヒストグラムより、1 日、7 日周期に含まれる BDA に注目し、各 BDA に対して周期性ヒストグラムを作成する。周期性ヒストグラムより最も投票が多かった時間帯を、各 BDA が一日の中で検出されやすい時間帯として抽出する。

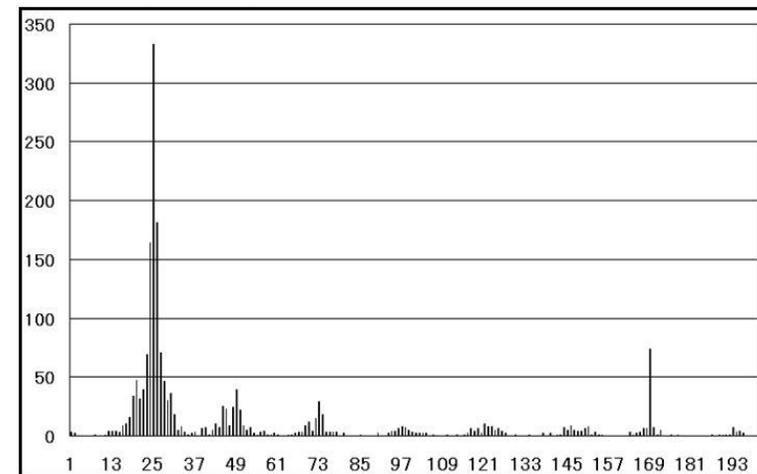


図 3. 周期性ヒストグラム

Step2. 共起ネットワークの構築と解析

抽出された BDA に対して、BDA を node、始めて検出された時期と検出されやすい時間帯の近さを edge としたネットワーク構造を作る。極端に距離が遠い edge を削除したネットワークに対してクラスタリング手法である CNM 法[10]を用いて時系列的に関係が密な BDA の組み合わせを解析する。図 4 は、実際にネットワークに対して、CNM 法を適応した例である。1 から 15 の BDA の組み合わせは、被験者の日々の出来事に関連していることが確認されている。クラスタに含まれる BDA を、何らかの出来

事に共起するものとして共起履歴に登録する。

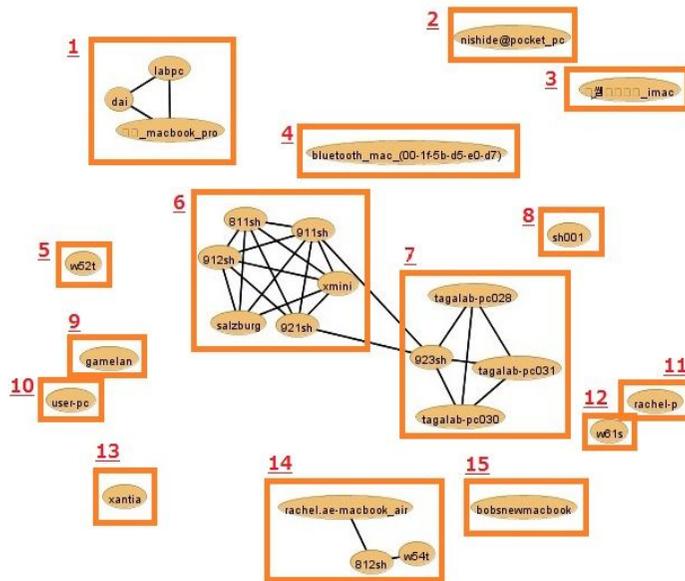


図 4. BDA の共起ネットワーク

4.2 教師データの作成

Step1. 修正履歴の蓄積

修正履歴が蓄積されるタイミングを以下に示す。

●複数のイベントタグが提示された場合

教師データを用いて日常検出した場合、教師データと類似した特徴が得られた時間帯には自動でイベントタグが付与されるが、付与されるイベントタグに複数の候補が挙がる場合がある。その際、ユーザに正解となるイベントタグを選択してもらう。複数のイベントタグが提示された原因としては、複数のイベントタグのサブセットになる BDA の組み合わせが得られたために、特徴の分離ができていないことが考えられる。そのため、選択と同時に提示されて時間帯に見られる BDA の組み合わせとイベントタグを新たに修正履歴に追加することで、学習による特徴の分離を試みる。また、提示された候補に正解が無い場合も、イベントタグを登録してもらい、提示された時間帯に見られる BDA の組み合わせと共に修正履歴に加える。

●ユーザに日常の正誤を求めた場合

自動でイベントタグが付与されなかった時間帯に対して、共起履歴を用いた日常検出をする。イベントタグが付与されなかった原因には、同じ出来事でも異なった BDA の組み合わせが存在する場合や、まだ検出されていない出来事が存在する場合が考えられる。そのため、イベントタグが付与されなかった時間帯に対して、日常検出をかけ正誤をユーザに問いかける。正しいと選択された場合、ユーザにイベントタグを設定してもらい BDA の組み合わせと共に修正履歴に加える。

●間違った非日常が提示された場合

教師データ、共起履歴を用いた日常検出によって判別できなかった時間帯に非日常があるとユーザにその正誤を問いかける。非日常でなかった場合、その時間帯に日常的な出来事があったと判断できるため、イベントタグの設定を求め、提示された時間帯に見られた BDA の共起と共に修正履歴に保持する。

Step2. 修正履歴を用いた教師データの学習

教師データの作成、及び更新には BDA の共起履歴の導出と同様にネットワーク解析を用いる。修正履歴に含まれる各イベントタグの修正項目より共起辞書を作成し、頻出する BDA の共起を基に共起ネットワークを作成する。共起辞書は、修正項目に含まれる BDA で総当りのペアを作ることで共起パターンを作成し、それぞれの共起が修正項目で生じる確率を調べることで作成する。図 5 の(1)は、イベントタグ A に対する修正項目から共起辞書を作成した例である。2, 3, 4, 5 は修正項目に現れる BDA を示しており、“|” で区切られた連番が一つの組み合わせである。このとき、2, 3, 4, 5 の総当りが共起として、各共起の修正項目の中での出現頻度を調べる。次に共起辞書より、頻出する BDA の共起を閾値判定によって抽出する。修正履歴には、イベントタグを判定するための最適な BDA 以外に、偶然参加していたなどの理由でノイズとなる BDA を含む組み合わせがある。そのため、共起頻度が低い特徴を削除することで最適な BDA の組み合わせに関連する共起のみを抽出する。図 5 の(2)は、共起辞書より閾値判定によって頻出する共起のみを抽出した例である。修正項目には、偶然その場に居合わせた BDA “2” が含まれているが、閾値判定によって除外されたのが分かる。この頻出する共起を用いて共起ネットワークを作成する。図 5 の(3)は、頻出する共起を用いて、ネットワーク構造を作成した例である。共起の BDA をノードにした無向グラフを生成し、無向グラフに対して CNM 法[10]を適応することで密な繋がりを示すクラスターを抽出する。そのクラスターに含まれる BDA をイベントタグに対する最適な組み合わせとして教師データに追加する。

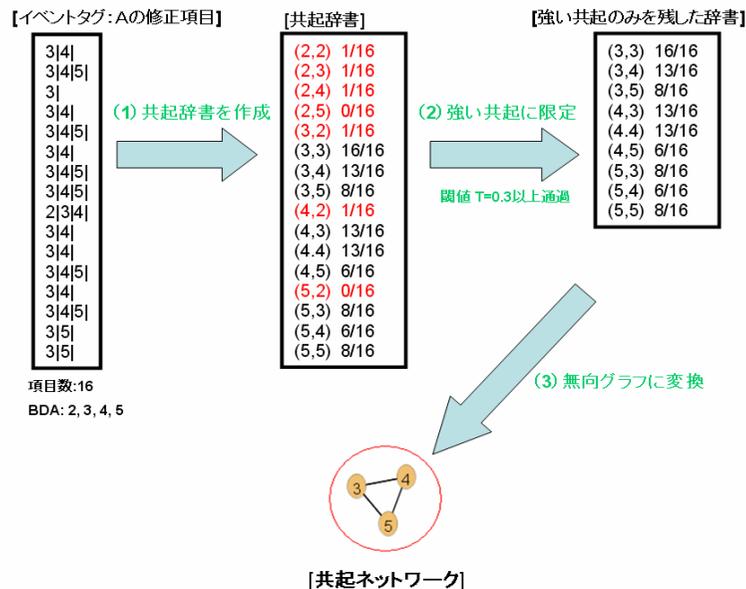


図 5. 教師データ更新の過程

4.3 日常の検出

教師データ、共起履歴はそれぞれ BDA の組み合わせで記述されている。それを用いて、入力されたデータの各時間帯の BDA の組み合わせと比較し、辞書に含まれる組み合わせが検出された時間帯を日常として検出する。以下に処理を示す。

データが取得された時刻: $n(n \in [1, N])$

Q_n : 時刻 n に取得された BDA の組み合わせ

C_m : 教師データ、共起履歴に記述された BDA の組み合わせ($m \in [1, M]$)

Th: 閾値

- Q_n と C_m に対して、(1)の計算をする。

$$\frac{Q_n \cup C_m}{Q_n \cap C_m} \dots (1)$$

IF (1) > Th THEN 時刻 n に対してラベル m を付与

以上の処理を入力された BT ログ $n(n \in [1, N])$ それぞれに対して繰り返す。

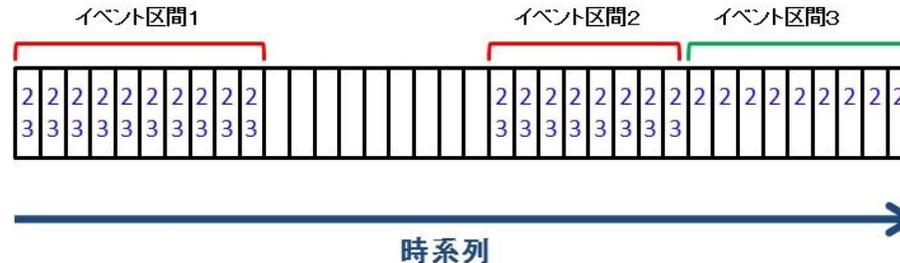


図 6. ラベル付与イメージ

(1) は、周期性をもつ BDA の組み合わせとの共通部分が多く、互いの集合のサイズが近ければ大きな値になる。これを用いて、周期性をもつ BDA の組み合わせに関連する時間帯にラベル付けをする。BDA の組み合わせにはバリエーションがあり、たとえ同じ出来事であっても組み合わせが安定することはない。そのため、生じる組み合わせを幅広く包含するように閾値は低く設定する。図 6 は、時系列順に並んだ BT ログの各 BDA の組み合わせに対してラベルを付与した例である。左から右に時系列順に並んでいる短冊は、分単位を表わしている。短冊の中に付与された数字は、検出された BDA の組み合わせに対して付与されたラベルを示す。前後のラベルの変化を捉え、出来事の変化を捉えている。この例のように、実際の場合にも複数のラベルが付与される例は多く存在し、複数のラベルが付与される場合は人が新たに加わるなど、実際のイベントに起きた変化を表している。そのため、前後のラベルとの変化を捉えることで同一のラベルが付与された区間を同一の出来事が起きている時間帯、すなわち日常として出力する。共起履歴を基にラベル付けされた場合は、始めて検出される出来事であるためユーザにイベントラベルの付与を求める。また、教師データによってラベルが付与された場合は、自動でイベントラベルを付与する。

5. 非日常の検出

非日常とは周期性を持たない出来事のことである。そのため、教師データ、共起履歴を用いた日常判定で、判定されなかった時間帯が非日常である可能性が高い。しかし、非日常は断定できるものではない。AirDiary が、日常の検出に失敗した可能性も考えられる。また、非日常の考え方も人それぞれである。そのため、AirDiary は非日常の推定と提案を行ない、非日常の判定はユーザに任せる。AirDiary は、非日常と推定される時間帯をユーザに提示し、その正誤を求める。正しいとされた場合、その時間帯

を非日常として出力する。また、ユーザにイベントタグを付与してもらうことで、日記に書くべき内容としてユーザに提示する。

6. テンプレート方式による短文形式の日記生成

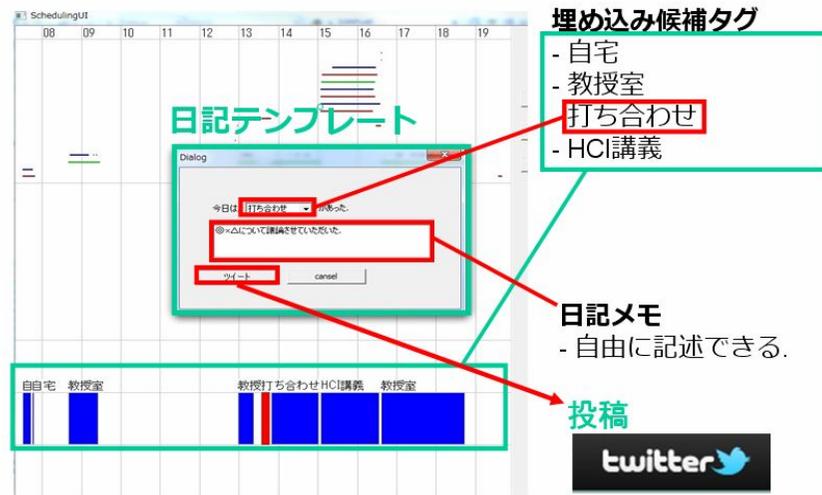


図 7. 日記テンプレート

日記の作成には、テンプレート方式をとる。図 7 に日記テンプレートを示す。あらかじめ用意されている日記テンプレートには、“今日は、「」があった。”といった形で、「」が存在する。「」に対して、入力した BT ログで検出されたイベントタグを候補としてユーザに提示し、イベントタグをユーザが選択することで日記を作成する。イベントタグは、ユーザによって任意に付けられるため、ユーザ個人のオリジナルの記述を残すことができるとともに、明示することを避けたタグを設定することでプライバシーな情報を守秘することができる。また、twitter と連携することで日記を公開、共有することができる。

7. 評価実験

7.1 実験方法

実験では、教師あり学習の有効性を示すために AirDiary に学習させながら BT ログを読み込ませた場合と、学習を完了し AirDiary に BT ログを読み込ませた場合の日常、非日常の検出精度を比較する。そのため、評価実験は 2 段階で行う。1 次実験では、AirDiary に対してデータセットを順次読み込ませ、教師データを学習しながら日常、非日常の判定を行うことで、常の検出精度、非日常の検出精度、複数のイベントタグが付与された確率を調べる。2 次実験では、1 次実験で作成した教師データを用いて学習させずデータセットに対して日常、非日常の判定を行い、同様の精度を測定する。1 次実験の結果と 2 次実験の結果を比較することで学習を加えた場合の精度の上昇率を比較する。データセットは、被験者 1 名の 2009 年 5 月から 2011 年 12 月末まで 2 年 1 ヶ月の BT ログを対象とし、2009 年 5 月から 2011 年 11 月末までのデータで共起履歴を作成する。学習データには、2011 年 12 月分のデータを用いる。正解セットは被験者が記録した行動履歴を基に作成した。

日常の検出精度は、数のイベントタグが付与された際の問いかけ、日常検出された際の問いかけの合計の数を全体の総数として正答率を算出する。日常の検出の正解と誤答は以下のように判断する。

●正解

- 日常が検出された際の問いかけに対して正答が得られた。(イベントタグの付与)
- 自動で正解であるイベントタグが付与された
- 複数のイベントタグが提示された際に正解タグが含まれた

●誤答

- 日常が検出された際の問いかけに対してイベントタグを付与しない
- 自動でイベントタグが付与された際に間違ったイベントタグが付与された
- 複数のイベントタグ候補が提示された際に正解が含まれない

非日常の検出精度は、非日常の正誤に対する問いかけの合計を全体数にして検出精度を測る。正答は、非日常の問いに対して正解の選択をしたとき、誤答は間違いとして修正を選択したときとする。複数のイベントタグが付与された確率は、教師データで日常検出がされた際に複数のイベントタグが付与されたため、ユーザに正答の選択を求める確率である。複数のイベントタグが付与された際の問の合計のうち、正解となるイベントラベルを含んでいた割合を調べる。複数のイベントタグが付与されることは、一度の検出で分離できない特徴の含有率を示すと同時に、日常が検出された際

にユーザに修正を求める回数を示すものである。以後これを、過剰問いかけ率と呼ぶ。

7.2 結果と考察

1次実験, 2次実験の結果を以下に示す。

表 1. 1次実験の結果

日常と判定された数	149
(うち誤検出の数)	21
非日常と判定された数	94
(うち誤検出の数)	34
複数のイベントタグが付与された数	57

日常の検出精度：85.9%
非日常の検出精度：63.8%
過剰問いかけ率：38.3%

表 2. 2次実験の結果

日常と判定された数	209
(うち誤検出の数)	10
非日常と判定された数	84
(うち誤検出の数)	25
複数のイベントタグが付与された数	108

日常の検出精度：95.2%
非日常の検出精度：70.2%
過剰問いかけ率：51.7%

表 1 は, 1次実験の結果を, 日常と判定された数, 日常と判定されたうちの誤検出の数, 非日常と判定された数, 非日常と判定されたうちの誤検出の数, 複数のイベントタグが付与された数で集計したものである。1次実験の結果は, 日常の検出精度は85.9%, 非日常の検出精度は63.8%, 過剰問いかけ率は38.3%となる。表 2 は, 2次実験の結果を日常と判定された数, 日常と判定されたうち誤検出の数, 非日常と判定された数, 非日常と判定されたうち誤検出の数, 複数イベントタグが付与された数を集計したものである。2次実験の結果は, 日常の検出精度は95.2%, 非日常の検出精度は70.2%, 過剰問いかけ率は51.7%となる。

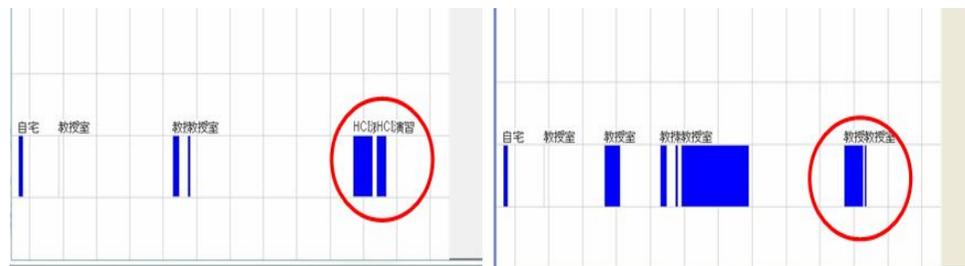


図 8. 自動イベントタグ割り当て精度の改善(左:1次実験 右:2次実験)

1次, 2次実験より提案する学習方式の有効性が示された。日常の検出精度は, 1次実験でも高い数値が得られ, 2次実験ではさらに高い精度が得られた。精度が向上した要因は, 誤り判定の減少である。1次実験の誤り判定の多くは, 教師データを用いた日常検出の際に, 自動で付与されたイベントタグが間違えるパターンである。その原因は, 複数のイベントが類似した BDA の組み合わせを特徴として持っているからである。そうした, 誤りの多くは修正履歴の蓄積が進み, 各イベントタグに最適な BDA の組み合わせが抽出されることで解決する。図 8 は, 1次実験, 2次実験で 2011 年 12 月 21 日のデータに対して, 教師データによる日常検出で自動にイベントタグが付与された結果である。左の図の丸が示す部分では, 間違ったイベントタグが付与され誤検出になったが, 右の図では誤検出が解消しただけでなく多くの時間帯に正しいイベントタグが付与された。修正履歴を蓄積することでイベントタグ「HCI 演習」と「教授室」を示す特徴を分離できた結果である。

非日常の検出精度は, 1次実験の結果では6割程度になり, 2次実験では7割まで上昇した。1次実験で蓄積した修正履歴が教師データに反映されたために, 非日常と判定された数と共に, 非日常と判定されたうちの誤判定の数が減少した結果である。誤って非日常と判定されるものの多くは, 大学の講義など7日周期を持つ出来事に集中している。大学の講義は, 多くの学生が参加していることから多くの BT 機器が検出される。しかし, 講義に欠かさず出席してくる学生は少ない, そのため同じ講義でも全く異なった BDA の組み合わせが生じている場合が多々あった。そのため, 1ヶ月では7日周期で生じる出来事を捉えるだけの修正履歴が蓄積できいことが推測される。

過剰問いかけ率は, 1次実験の結果では4割程度の頻出度だが, 2次実験では5割程度に上昇した。過剰問いかけ率は, 教師データで日常検出がされた際に同じ時間帯に複数のイベントタグが付与されたため, ユーザに正答の選択を求める確率である。したがって, 確率が低いほどユーザの負担は少なく, かつ自動でイベントタグが付与される出来事の特徴が分離できていることが考えられる。図 9 は, 実際に複数のイベントタグが提示された例である。この提示がされた時間帯では, 1つの BDA が取得された。それに対して, 提示されたイベントタグは5つである。これは, 5つのイベントタグを示す BDA の組み合わせの要素に, ただ1つ検出された BDA が含まれているためである。そのため, イベントタグを自動で決定できなかった。しかし, これは AirDiary の処理が問題ではなく, 取得された BT ログの濃度の問題である。常時複数の BDA が取得されるような状況であるならば, 他のイベントタグと分離するだけの BDA の組み合わせの変化が得られる可能性がある。しかし, この場合は他と差別化するだけの BDA が得られていない。これは, 常時多くの BT 機器の取得が難しい社会の状況による影響が大きい。



図9. 複数のイベントタグが提示された例

8. おわりに

本研究では、BT ログを用いた半自動日記作成ツールである AirDiary を作成した。ユーザがモバイル端末によって常時取得した BT ログを入力に、日々の生活の中に隠れた非日常的な出来事をユーザと協調的に判定する。判定された出来事をテンプレートに埋め込むことで短文形式の日記の作成を支援する。また、twitter と連携することで日記の公開、共有が可能である。日常の判定には、長期の蓄積がある BT ログを解析することで導出した共起履歴と、ユーザの修正履歴を用いて導かれる教師データを用いる。AirDiary は、日常、非日常が検出されるたびに正誤を問い、誤検出した場合のユーザの修正履歴を学習することで、一度検出した出来事を自動で検出する。提案した学習方式で AirDiary の日常の検出精度 95.2%、非日常の検出精度 70.2%と高い精度が得られた。

今後は、日記作成支援を重視して研究を進める。現在、日記テンプレートは 1 種類のみだが、複数のイベントタグに対応するなどテンプレートのパターンを追求する必要がある。それによって、日記の文章表現を追求するものである。また、日記といっても文章のみで構成されているわけではない。そのため、写真管理ソフトとの連携やペイントツールとの連携も視野に研究を進める。また、学習による過剰問い合わせ率の

上昇も解決することが必要である。過剰問い合わせ率は、学習によって導出された BDA の組み合わせが複数のイベントのサブセットとなるときに起きる。学習によって、各出来事にとって代表となる BDA の組み合わせが抽出されていることには間違いはない。事実、多くの出来事は分離可能である。しかし、出来事の種類によっては、代表となる BDA の組み合わせよりもそれに付随する BDA が重要となる場合がある。そうした、出来事の種類を考慮した学習手法を開発する必要がある。

謝辞

的確な助言をいただいた京都大学情報学研究科の中村聡史准教授に感謝する。

参考文献

- [1] mixi. <http://mixi.jp/>.
- [2] Facebook. <http://ja-jp.facebook.com/>.
- [3] twitter. <http://twitter.com/>.
- [4] 角康之, 坂本竜基, 中尾恵子, 間瀬健二. “コミックダイアリ : 経験や興味を伝え合うための漫画日記.” インタラクシオン2002, pp. 101-108. 情報処理学会, 2002.
- [5] キセキ. <http://lifelog.machi.goo.ne.jp/>.
- [6] T.Nicolai, and H.Kenn. “Towards Detecting Social Situations with Bluetooth”. Adjunct Proc. Ubicomp 2006, Sep, 2006.
- [7] V.Kostakos, T.Nicolai, E.Yoneki, E.O’Neill, H.Kenn, J.Crowcroft. “Understanding and measuring the urban pervasive infrastructure”. Personal and Ubiquitous Computing, Springer, Vol.13, No.5, pp.355-364, 2009.
- [8] H.Kim, K.Cho. “BlueSense: Sensing Blue Whales”. <http://urban.cens.ucla.edu/cs219/images/0/0b/BlueSense.pdf>.
- [9] 牛越達也, 河野恭之. “Bluetoothデバイス検出履歴の長期的観測に基づく周期性の検証”. WISS2009. pp159-160. 2009.
- [10] A.Clauset, M.E.J.Newman and C.Moore. “Finding community structure in very large networks.” *Phy. Rev. E*, 70, 0066111, 2004.